

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 13:46:37 ; Search time 14114 Seconds
(without alignments)
17476.634 Million cell updates/sec

Title: US-09-964-956-12
Perfect score: 5691
Sequence: 1 atgaagccatgcctcgaa.....gcttagacagctgaataaa 5691

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hgt_hum:*
31: em_hgt_in:*
32: em_hgt_inv:*
33: em_hgt_other:*
34: em_hgt_mus:*
35: em_hgt_pln:*
36: em_hgt_rod:*
37: em_hgt_mam:*
38: em_hgt_vrt:*
39: em_sy:*
40: em_hgt_hum:*
41: em_hgt_mus:*
42: em_hgt_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5691	100.0	5691	6	AX503782
2	5680.6	99.8	6367	6	AX429828
3	4714.2	82.8	7022	10	AB073228
4	2939.6	51.7	6387	6	AX405895
5	2938	51.6	6263	9	AB007932
6	2778.4	48.8	6730	6	AX305728
7	2778.4	48.8	6730	10	D86949
8	2659.4	46.7	6147	6	AX704766
9	2578.4	45.3	6466	10	D86948
10	2535.4	44.6	6039	9	HSSEXGENE
11	2533.8	44.5	5895	6	AX686475
12	2423.2	42.6	5263	9	HSNOVPROT
13	2387.8	42.0	6893	10	AX122289
14	2366.8	41.6	6163	5	XELPLEX
15	2312.8	40.6	5982	10	D86950
16	2240.2	39.4	5986	9	AB046770
17	2146.8	37.7	3985	9	HSOCTPROT
18	2015.6	35.4	6693	6	BD183298
19	1554	27.3	2875	9	BC009343
20	1377	24.2	2015	9	BC028744
21	1371.8	24.1	3310	9	AK123428
22	1188.4	20.9	179145	9	AC009785
23	1185.2	20.8	188090	9	AC011625
24	1177.4	20.7	179592	2	AC009723
25	1167.4	20.5	2284	10	BC056475
26	1157.8	20.3	3666	6	AX833678
27	1157.8	20.3	3666	6	AX095606
28	1026.6	18.0	3556	9	HSMB05617
29	930.8	16.4	221370	2	AC126205
30	930.8	16.4	228898	2	AC131554
31	930.8	16.4	228922	2	AC101241
32	915.2	16.1	1817	9	BC006193
33	878.2	15.4	4445	9	AK126101
34	793.8	13.9	2597	6	AX574564
35	793.8	13.9	2597	9	AY358850
36	759.8	13.4	1029	9	HSMB02037
37	673.6	11.8	776	4	AF176418
38	630.8	11.1	3560	6	AX747045
39	630.8	11.1	3560	6	AK091538
40	625	11.0	74655	2	AC025595
41	613.2	10.8	1403	10	BC024509
42	598.4	10.5	56201	2	AC144970
43	591.4	10.4	601	6	AX088116
44	586.6	10.3	2414	6	BD127385
45	586.6	10.3	2414	9	AK074960

ALIGNMENTS

RESULT 1	AX503782	5691 bp	DNA	linear	PAT 27-SEP-2002
LOCUS	AX503782				
DEFINITION	Sequence 12 from Patent WO0226826.				
ACCESSION	AX503782				
VERSION	AX503782.1	GI:23385966			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
REFERENCE					
AUTHORS	Gerlach, V.L., MacDougall, J.R., Smithson, G., Miller, I., Stone, D., Gutierrez, E., Ellerman, K., Grose, W.M., Alsbrook, P., Lepley, D.M., Burgess, C.E., Padigara, M., Kekuda, R., Szytek, K.A., Leach, M.D. and				

Shinkets, R.A.
Proteins and nucleic acids encoding same
Patent: WO 026826-A 12 04-APR-2002;
JOURNAL Curagen Corporation (US)

FEATURES
Source
1. 5691
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 5691; DB 6; Length 5691;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5691; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGAAAGCCATGCCCCTGAACTGACCTGCTTCTCTCCCACTCTCATGTGGCATG 60
1 ATGAAAGCCATGCCCCTGAACTGACCTGCTTCTCTCCCACTCTCATGTGGCATG 60
61 GGCTCTTCCACTTGTCTACCCCGACCCCGCCGCTGTCAGAAAGCAGCGGTATT 120
61 GGCTCTTCCACTTGTCTACCCCGACCCCGCCGCTGTCAGAAAGCAGCGGTATT 120
121 GTCAATTCGAGAGAGAGCCCGCAGAGGTTTCAATCCTGTGTGTGATGAGAGACA 180
121 GTCAATTCGAGAGAGAGCCCGCAGAGGTTTCAATCCTGTGTGTGATGAGAGACA 180
181 GAGACATTTTACTTGGGGGCGCTCATGCGATTTCAGAGCTTTCAGAGCCTGAAAGTTC 240
181 GAGACATTTTACTTGGGGGCGCTCATGCGATTTCAGAGCTTTCAGAGCCTGAAAGTTC 240
241 TTGGTGAAGCATGAGAGAGAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
241 TTGGTGAAGCATGAGAGAGAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
301 GTCCAGACCTGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
301 GTCCAGACCTGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
361 GACTCAAG 420
361 GACTCAAG 420
421 CTGAGGCTGAG 480
421 CTGAGGCTGAG 480
481 TCAGGTTCACAG 540
481 TCAGGTTCACAG 540
541 GACAAGCTGTTCAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
541 GACAAGCTGTTCAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
601 AGCCGAG 660
601 AGCCGAG 660
661 GAGTTCGTGGCTGATGATTAAGATCCCTTCGAGACCTTCACATCATCCTGACTTT 720
661 GAGTTCGTGGCTGATGATTAAGATCCCTTCGAGACCTTCACATCATCCTGACTTT 720
721 GATATCTATATGCTATGCTTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
721 GATATCTATATGCTATGCTTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
781 CCTGAG 840
781 CCTGAG 840
841 GTGAGGCTTTGAG 900
841 GTGAGGCTTTGAG 900

841 GTGAGGCTTTGAG 900
901 GAGGCGAGTGGAG 960
901 GAGGCGAGTGGAG 960
961 GTGCTTGGCAG 1020
961 GTGCTTGGCAG 1020
1021 AAGGCGCAG 1080
1021 AAGGCGCAG 1080
1081 AAGCAGATTAATGAG 1140
1081 AAGCAGATTAATGAG 1140
1141 CTGAGACCTGGCTGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
1141 CTGAGACCTGGCTGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
1201 GAGCATTAATCTTGTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
1201 GAGCATTAATCTTGTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
1261 GGAATTCCTGCTTTCAG 1320
1261 GGAATTCCTGCTTTCAG 1320
1321 AAGAACACTCTCTGAG 1380
1321 AAGAACACTCTCTGAG 1380
1381 GATGAGACCAAG 1440
1381 GATGAGACCAAG 1440
1441 GTCTCCGAGATTAATGCTTCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
1441 GTCTCCGAGATTAATGCTTCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
1501 GAGCTCAAG 1560
1501 GAGCTCAAG 1560
1561 GAGCTCAAG 1620
1561 GAGCTCAAG 1620
1621 CGGTGAG 1680
1621 CGGTGAG 1680
1681 CTGAGAGTCCATCCCAATATATCTGCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
1681 CTGAGAGTCCATCCCAATATATCTGCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
1741 AGCTAATGATCCCGAG 1800
1741 AGCTAATGATCCCGAG 1800
1801 ATGATGAGGCTGAGTCTGAG 1860
1801 ATGATGAGGCTGAGTCTGAG 1860
1861 CCGCGATATATCAAG 1920
1861 CCGCGATATATCAAG 1920
1921 GAGACCGGAG 1980
1921 GAGACCGGAG 1980

```
QY 1981 TCGTGGCTGTCTTCCGTGGAGATCCATACCGCTGCCAATGCTGAATATACCGGCAATGTC 2040
Db 1981 TCGTGGCTGTCTTCCGTGGAGATCCATACCGCTGCCAATGCTGAATATACCGGCAATGTC 2040
QY 2041 TGCACCAATGACCCCAAGACCTGTCTTCCAGAGAGAGCGAGTGAAGTGTCCCGAGAGAC 2100
Db 2041 TGCACCAATGACCCCAAGACCTGTCTTCCAGAGAGAGCGAGTGAAGTGTCCCGAGAGAC 2100
QY 2101 TCGCCCAAGCTGTCTCCGATGGAACAAGATCTGTGTCCCTGTGAGAGTGAATCAAGCTTATC 2160
Db 2101 TCGCCCAAGCTGTCTCCGATGGAACAAGATCTGTGTCCCTGTGAGAGTGAATCAAGCTTATC 2160
QY 2161 ACGCTGAAGGCCAAGAACCTCCCGCAGGCCAGTCTGTGGGCAAGCTGTGCTACGAATGATC 2220
Db 2161 ACGCTGAAGGCCAAGAACCTCCCGCAGGCCAGTCTGTGGGCAAGCTGTGCTACGAATGATC 2220
QY 2221 CTCAACATTCAGGGGAGGAGAGAGAGTGTCCCGCTGTGAGTCAAGCTGTGAGTCAAGCTG 2280
Db 2221 CTCAACATTCAGGGGAGGAGAGAGTGTCCCGCTGTGAGTCAAGCTGTGAGTCAAGCTG 2280
QY 2281 CAGTGGCCAGAACACCTCTTATTTCCATGAGAGAGTGAATCAACAACCTGTGAGTCAAG 2340
Db 2281 CAGTGGCCAGAACACCTCTTATTTCCATGAGAGAGTGAATCAACAACCTGTGAGTCAAG 2340
QY 2341 TTGACAGTGTGTGGAATGGGCACTTCAACATGACAAACCCAGCTCAAGATTAAGTTTAC 2400
Db 2341 TTGACAGTGTGTGGAATGGGCACTTCAACATGACAAACCCAGCTCAAGATTAAGTTTAC 2400
QY 2401 CTCTAACAGTGTGAGAGAGTGTGTGAGAGTGTGGGCTGTGAGTCAAGAGCTGACCCAGAC 2460
Db 2401 CTCTAACAGTGTGAGAGAGTGTGTGAGAGTGTGGGCTGTGAGTCAAGAGCTGACCCAGAC 2460
QY 2461 TTGCGATGTGTGTGTGCGCAGAGGCGCAGGCGCAGTGTGACCTGTGCGCAGACCTGTGCGC 2520
Db 2461 TTGCGATGTGTGTGTGCGCAGAGGCGCAGGCGCAGTGTGACCTGTGCGCAGACCTGTGCGC 2520
QY 2521 CAGGAGAGCCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
Db 2521 CAGGAGAGCCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
QY 2581 ACAGAGATTAATCCCGGTGACAGAGCGCCCGGAGAGGGGAGCCAAAGTCACTATCCAGAGG 2640
Db 2581 ACAGAGATTAATCCCGGTGACAGAGCGCCCGGAGAGGGGAGCCAAAGTCACTATCCAGAGG 2640
QY 2641 GAGAACTGTGGGCTGTGAAATTTTGGGCAATGCTGTCCATGTCAAGTGTGTGTGTGTGTGT 2700
Db 2641 GAGAACTGTGGGCTGTGAAATTTTGGGCAATGCTGTCCATGTCAAGTGTGTGTGTGTGTGT 2700
QY 2701 TGCACGCTTTTATGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2760
Db 2701 TGCACGCTTTTATGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2760
QY 2761 GCCAAGCCCAAGCAGATGCAAGGCTTCCGAGAGATCTGTGTGTGTGTGTGTGTGTGTGTAA 2820
Db 2761 GCCAAGCCCAAGCAGATGCAAGGCTTCCGAGAGATCTGTGTGTGTGTGTGTGTGTGTGTAA 2820
QY 2821 TTGATAGGCGCGGTCTCTCAAGTCAATTAATTTATGATGATGATGATGATGATGATGATGATG 2880
Db 2821 TTGATAGGCGCGGTCTCTCAAGTCAATTAATTTATGATGATGATGATGATGATGATGATGATG 2880
QY 2881 CCCAGCGGGGAGCCCAATGTCCGAGAGAGCCCAAGTGAACCATCAAGGAGCCAACTGTAT 2940
Db 2881 CCCAGCGGGGAGCCCAATGTCCGAGAGAGCCCAAGTGAACCATCAAGGAGCCAACTGTAT 2940
QY 2941 GCCGGAAGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3000
Db 2941 GCCGGAAGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3000
QY 3001 CCATCTTACATTTGTGTGACAAACCAATCTGATGATGATGATGATGATGATGATGATGATGATG 3060
Db 3001 CCATCTTACATTTGTGTGACAAACCAATCTGATGATGATGATGATGATGATGATGATGATGATG 3060
QY 3061 GTGCAAGTGTGACAGGAGCCAAAGATCCACAGAGACCTGTGTCTTCAATGATGTGAAAGACCCC 3120
Db 3061 GTGCAAGTGTGACAGGAGCCAAAGATCCACAGAGACCTGTGTCTTCAATGATGTGAAAGACCCC 3120
QY 3121 ACCATCTGTGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3180
Db 3121 ACCATCTGTGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3180
QY 3181 TGGGGAGACCCACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3240
Db 3181 TGGGGAGACCCACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3240
QY 3241 GAGCAGATTAATTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3300
Db 3241 GAGCAGATTAATTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3300
QY 3301 CTGCTGTGGTCTGTGACCAAGCAGTCAAGCTGACCGAGAGGCCAGAGAGTTTGTGCTTC 3360
Db 3301 CTGCTGTGGTCTGTGACCAAGCAGTCAAGCTGACCGAGAGGCCAGAGAGTTTGTGCTTC 3360
QY 3361 ATCTGTGACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3420
Db 3361 ATCTGTGACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3420
QY 3421 AACCGGTGTGTGAGGCTTTGTGCTGTGAGGATCTGTGAGATCTGTGAGCTGAGCTGTGAGCTGAG 3480
Db 3421 AACCGGTGTGTGAGGCTTTGTGCTGTGAGGATCTGTGAGATCTGTGAGCTGAGCTGTGAGCTGAG 3480
QY 3481 ATCATCTTAAAGGAGCAAGAACCTGTATCCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3540
Db 3481 ATCATCTTAAAGGAGCAAGAACCTGTATCCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3540
QY 3541 TACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3600
Db 3541 TACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3600
QY 3601 TGGGATCTCCCAACCTCATGTGAGAGGCAAAAGTATGAGCCGTGTGTGTGTGTGTGTGTGTGTGT 3660
Db 3601 TGGGATCTCCCAACCTCATGTGAGAGGCAAAAGTATGAGCCGTGTGTGTGTGTGTGTGTGTGTGT 3660
QY 3661 TACTCCCCGGGATGTGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3720
Db 3661 TACTCCCCGGGATGTGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3720
QY 3721 AGCATGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3780
Db 3721 AGCATGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3780
QY 3781 AAAAGCAAGTCCCGGAAAGTGAACCTCAAGCTGTGAAGCGGCTGACATGATGATGATGATGATGATG 3840
Db 3781 AAAAGCAAGTCCCGGAAAGTGAACCTCAAGCTGTGAAGCGGCTGACATGATGATGATGATGATGATG 3840
QY 3841 CTGAGATCCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3900
Db 3841 CTGAGATCCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3900
QY 3901 CATGAGTGAACAGTGAACCTGTGATGAGCCGGGATTCGTTCTGTGATGATGATGATGATGATGATGATG 3960
Db 3901 CATGAGTGAACAGTGAACCTGTGATGAGCCGGGATTCGTTCTGTGATGATGATGATGATGATGATGATG 3960
QY 3961 ACCATGTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4020
Db 3961 ACCATGTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4020
QY 4021 CCGGGCTACCGGAGAGAGCTGTGTGAGAAAGGCTGTGAGCTCTTCCCGCAGCTCATGAC 4080
Db 4021 CCGGGCTACCGGAGAGAGCTGTGTGAGAAAGGCTGTGAGCTCTTCCCGCAGCTCATGAC 4080
QY 4081 AACAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4140
Db 4081 AACAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4140
QY 4141 CGGACCGGTGTGCAAGCTGTGCTCATCATCATGATGATGATGATGATGATGATGATGATGATGATGATG 4200
```

Db	4141	CGGACCGTGGCAACGTGGCTCACTCATCATGACCGTGGTGGCAAGCAAGCTGGAGTAC	4200
Qy	4201	GGCAGTGAATGTGTAAAGCAAGTGTGTGGCCGACCTCATTTGACAAAGAACTTGGAGACGAC	4260
Db	4201	GGCAGTGAATGTGTAAAGCAAGTGTGTGGCCGACCTCATTTGACAAAGAACTTGGAGACGAC	4260
Qy	4261	AACCACTTAAGCTGTGTCTGTCAGAGAGACTGAGTCAAGTGGCTGAGAAAGATGTGCACAAAT	4320
Db	4261	AACCACTTAAGCTGTGTCTGTCAGAGAGACTGAGTCAAGTGGCTGAGAAAGATGTGCACAAAT	4320
Qy	4321	TGGTTTAATTTCCTCCTCTTCAAAAGTTCCCTCAAGAAAGTGTGTGGAGAGCCCTCTTCTCC	4380
Db	4321	TGGTTTAATTTCCTCCTCTTCAAAAGTTCCCTCAAGAAAGTGTGTGGAGAGCCCTCTTCTCC	4380
Qy	4381	CTGTCTGTGGCCATCAAGCAGACATGGAGAAAGGCCCCCATTTGACCCATCACTGGGCGAG	4440
Db	4381	CTGTCTGTGTGGCCATCAAGCAGACATGGAGAAAGGCCCCCATTTGACCCATCACTGGGCGAG	4440
Qy	4441	GCCCCGTACTCTCTTGAAGGAGACAAAGTCAATCCGCGACAGATTTGACTTACAAAACTCTG	4500
Db	4441	GCCCCGTACTCTCTTGAAGGAGACAAAGTCAATCCGCGACAGATTTGACTTACAAAACTCTG	4500
Qy	4501	GTCTTAGCTGTGTGACGCCACAGACAAATCCAAAGCCCCGAGGTCCCATTTAAAGATCTCT	4560
Db	4501	GTCTTAGCTGTGTGACGCCACAGACAAATCCAAAGCCCCGAGGTCCCATTTAAAGATCTCT	4560
Qy	4561	AATGTGACACACATCACTCAAGTCAAGAGAAAGTTTGGATATGCCATCTTCAAGAAATGTG	4620
Db	4561	AATGTGACACACATCACTCAAGTCAAGAGAAAGTTTGGATATGCCATCTTCAAGAAATGTG	4620
Qy	4621	CCTTGTCCACCGGCCCCAAAGCTGCAGATATGATCTGAGTGGGCGACAAAGAAATGGG	4680
Db	4621	CCTTGTCCACCGGCCCCAAAGCTGCAGATATGATCTGAGTGGGCGACAAAGAAATGGG	4680
Qy	4681	GCAAGATATCTTGTGAGATGGAAGACATCAACCAAGATTGAAGATATTTGAAGAGCGA	4740
Db	4681	GCAAGATATCTTGTGAGATGGAAGACATCAACCAAGATTGAAGATATTTGAAGAGCGA	4740
Qy	4741	CTGACACACTGGCCCACTACAGGTGGCAGATGGTTCCTGTGTGGCATTTAGTGTCCAG	4800
Db	4741	CTGACACACTGGCCCACTACAGGTGGCAGATGGTTCCTGTGTGGCATTTAGTGTCCAG	4800
Qy	4801	CAGGTGACGCGCTATTAACGAGTGAACAATCCACCGTTCACAGACCTCAGACAGTAA	4860
Db	4801	CAGGTGACGCGCTATTAACGAGTGAACAATCCACCGTTCACAGACCTCAGACAGTAA	4860
Qy	4861	TATGAAACATGATCCGATACAGCGGCGAGCCCCCAGACGCTCCGCTCACGACACTATG	4920
Db	4861	TATGAAACATGATCCGATACAGCGGCGAGCCCCCAGACGCTCCGCTCACGACACTATG	4920
Qy	4921	ATCACTTCGACCTGAGAGTGGAGTCAAGATGTGGACCTATGTGAAGAACCAACAGACAC	4980
Db	4921	ATCACTTCGACCTGAGAGTGGAGTCAAGATGTGGACCTATGTGTGAAGAACCAACAGACAC	4980
Qy	4981	GGAAGACAGAAGAGGGGAGCCGGGGAGCAAGATGGTGTCTGAATTTCACTGACCCGA	5040
Db	4981	GGAAGACAGAAGAGGGGAGCCGGGGAGCAAGATGGTGTCTGAATTTCACTGACCCGA	5040
Qy	5041	CTCTGTGGCCATTAAGGCGACACTGTGAAAGATTTGTGATACCTCTTTAGACCAATCTTC	5100
Db	5041	CTCTGTGGCCATTAAGGCGACACTGTGAAAGATTTGTGATACCTCTTTAGACCAATCTTC	5100
Qy	5101	AGCAGCGCACACCGTGGCTTGTGCTTGCCTTGCCTGGCCATCAAGTACAATGTTTGAATCTCTG	5160
Db	5101	AGCAGCGCACACCGTGGCTTGTGCTTGCCTTGCCTGGCCATCAAGTACAATGTTTGAATCTCTG	5160
Qy	5161	GATGACAGAGCTGATTAACATGGCAATTAATGATCCCGACGATCCGCAATCTCTGGAAGAC	5220
Db	5161	GATGACAGAGCTGATTAACATGGCAATTAATGATCCCGACGATCCGCAATCTCTGGAAGAC	5220
Qy	5221	AATTGCTGTCCCTTGAAGTTTGGGTCAACATGATCAAGAACCCGCAATTTGTGTTTAC	5280

Db	5221	AATTGCGTGGCCCTGGAGGTTTGGGTGCAATGATGTCAAAGAACCCGCGAGTTGTGTGGAC	5286
QY	5281	ATTCATATAGAACACACATACAGAGAGCGTGGCTCTCTGTGTGGGTGACACTTCAATGAC	5340
Db	5281	ATCCATATAGAACACACATACAGAGAGCGTGGCTCTCTGTGTGGGTGACACTTCAATGAC	5340
QY	5341	TCTTGCTCCAGCTGACGACACCCGGCTGGGACAGGACTCGCCCTCCAAACAGCTGCTGAT	5400
Db	5341	TCTTGCTCCAGCTGACGACACCCGGCTGGGACAGGACTCGCCCTCCAAACAGCTGCTGAT	5400
QY	5401	GCCAAAGACATATCCCAAGTACAAATAATGGGTGAGAGGTATATCTGACATATAGAGGAG	5460
Db	5401	GCCAAAGACATATCCCAAGTACAAATAATGGGTGAGAGGTATATCTGACATATAGAGGAG	5460
QY	5461	ATGCGCAGCCCATACACGACCAAGACATGATAGCATATCTGCGCTGAGCAGCTGCCGATGAC	5520
Db	5461	ATGCGCAGCCCATACACGACCAAGACATGATAGCATATCTGCGCTGAGCAGCTGCCGATGAC	5520
QY	5521	ATGATATAGTTCACACACCATGAGAGGCACTCTGACAGATCTTTTCCCTATGTGGGCAATAC	5580
Db	5521	ATGATATAGTTCACACACCATGAGAGGCACTCTGACAGATCTTTTCCCTATGTGGGCAATAC	5580
QY	5581	AGCGAGGAGATTCCTTGGACCTCTGACCAACAGATGACCCAGTGTGGGAGACAGAAACTGGCC	5640
Db	5581	AGCGAGGAGATTCCTTGGACCTCTGACCAACAGATGACCCAGTGTGGGAGACAGAAACTGGCC	5640
QY	5641	TACAAACTATAGAACAGTCATPACCCCTCATAGCTTACACAGCTGAATATAA	5691
Db	5641	TACAAACTATAGAACAGTCATPACCCCTCATAGCTTACACAGCTGAATATAA	5691
RESULT 2			
LOCUS	AX492982	6367 bp	DNA
DEFINITION	Sequence 19 from Patent WO02059312.	Linear	PAT 26-SEP-2002
ACCESSION	AX492982		
VERSION	AX492982.1	GI:23338637	
KEYWORDS			
ORGANISM	Homo sapiens (human)		
AUTHORS	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	Kallik,D.A., Lee,S., Xu,Y., Yao,M.G., Yue,H., Bandman,O.B., Butford,N., Gandhi,A.R., Grall,R.C., Lal,R.G., Lu,D.A., Lu,Y., Tang,T.Y., Dugan,B.M., Gletten,K.J., Hillman,J.T., Honcheil,C.D., Ramkumar,J., Wallis,N.K. and Warren,B.A. Cell adhesion proteins Patent: WO 02059312-A 19 01-AUG-2002; JOURNAL INCYTE GENOMICS INC (US)		
FEATURES	location/Qualifiers		
source	1..6367 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606" /note="Incyte ID No: 7156379CB1"		
ORIGIN			
Query Match	99.8%; Score 5680.6; DB 6; Length 6367;		
Best Local Similarity	99.9%; Pred. No. 0;		
Matches 5683; Conservative	0; Mismatches 4; Indels 0; Gaps 0;		
QY	1	ATGAAAGCCATGCGCTTGGAACTGGACCTGCTTCTCTCCACCTCTCATGTGGGCGATG	60
Db	566	ATGAAAGCCATGCGCTTGGAACTGGACCTGCTTCTCTCCACCTCTCATGTGGGCGATG	625
QY	61	GGCTCTCCACTTGTGTCACCCGGCAGCAGCCCGGTGTCGCCAGAAAGCAGCGTCAATT	120
Db	626	GGCTCTCCACTTGTGTCACCCGGCAGCAGCCCGGTGTCGCCAGAAAGCAGCGTCAATT	685
QY	121	GTCACTTCCGAGAGAGCCCGCCGAGGAGTTTCAATCACTGTTGTGATGAGAGACA	180
Db	686	GTCACTTCCGAGAGAGCCCGCCGAGGAGTTTCAATCACTGTTGTGATGAGAGACA	745

181 GGACACATTTACTTGGGGGCGCTGATCGGATTTTACAGCTTCCAGCGACCTGAGGTC 240
Db GGACACATTTACTTGGGGGCGCTGATCGGATTTTACAGCTTCCAGCGACCTGAGGTC 805
Oy TTGGTGAACGATGAGACAGGGCCGAGAGACAAACCCCAAGTTTACCCAGCCCGATC 300
Db TTGGTGAACGATGAGACAGGGCCGAGAGACAAACCCCAAGTTTACCCAGCCCGATC 865
Oy GTCCAGACCTGCAATGAGCCCTTCAACCAACATGTCAACAGATGCTCTCTATA 360
Db GTCCAGACCTGCAATGAGCCCTTCAACCAACATGTCAACAGATGCTCTCTATA 925
Oy GACTACAAAGAGAAACAGGCTGATGCGCTGAGAGCTGATACCAAGGATCTGCAAGCTG 420
Db GACTACAAAGAGAAACAGGCTGATGCGCTGAGAGCTGATACCAAGGATCTGCAAGCTG 985
Oy CTGAGGCTGAGAGACCTCTTCAAGCTGAGGAGGCTTATCATTAAGAGAGCACTATCTG 480
Db CTGAGGCTGAGAGACCTCTTCAAGCTGAGGAGGCTTATCATTAAGAGAGCACTATCTG 1045
Oy TCAGGTGTCACAGAGAGGAGGCTGAGTGGATGCTGCTGATGCTGCTGATGCTGCTGAT 540
Db TCAGGTGTCACAGAGAGGAGGCTGAGTGGATGCTGCTGATGCTGCTGATGCTGCTGAT 1105
Oy GACAGAGCTGATCATTTGCCACGAGAGTGGATGAGAGCCGAGATATTTTCCACCATCTCC 600
Db GACAGAGCTGATCATTTGCCACGAGAGTGGATGAGAGCCGAGATATTTTCCACCATCTCC 1165
Oy AGCCGGAACCTGACCAAGAACTCTGAGGCGGATGGCACTGCTGCTGCTGCTGCTGAT 660
Db AGCCGGAACCTGACCAAGAACTCTGAGGCGGATGGCACTGCTGCTGCTGCTGCTGAT 1225
Oy GAGTTGCTGAGCTGATGATTAAGATCCCTTCCGACACCTTCAACCATCATCTGATCTT 720
Db GAGTTGCTGAGCTGATGATTAAGATCCCTTCCGACACCTTCAACCATCATCTGATCTT 1285
Oy GATATCTATCTATGCTATGCTTTTAAAGAGTGGCACTTGTCTACTTTTGAACCTTCAA 780
Db GATATCTATCTATGCTATGCTTTTAAAGAGTGGCACTTGTCTACTTTTGAACCTTCAA 1345
Oy GATATCTATCTATGCTATGCTTTTAAAGAGTGGCACTTGTCTACTTTTGAACCTTCAA 1345
Db GATATCTATCTATGCTATGCTTTTAAAGAGTGGCACTTGTCTACTTTTGAACCTTCAA 1405
Oy CCTGAGATGCTGCTCCACAGGCTCCACCAAGAGAGAGTGTATATCAATCCAGCTC 840
Db CCTGAGATGCTGCTCCACAGGCTCCACCAAGAGAGAGTGTATATCAATCCAGCTC 1405
Oy GTGAGGCTTTGCAAGAGAGAGACAGGCTTCACTCTATGTAAGAGTGGCCATTTGGCTGT 900
Db GTGAGGCTTTGCAAGAGAGAGACAGGCTTCACTCTATGTAAGAGTGGCCATTTGGCTGT 1465
Oy GAGGCGCATGAGGCTGAGATGACCGGCTGCTGAGAGGCTGCTCACTGCTCAAGCGAGG 960
Db GAGGCGCATGAGGCTGAGATGACCGGCTGCTGAGAGGCTGCTCACTGCTCAAGCGAGG 1525
Oy GTGCTTGGCAGGACCTTGAAGTCCATCAAGATGATGATGCTGCTTCAACCGTCTTCTCC 1020
Db GTGCTTGGCAGGACCTTGAAGTCCATCAAGATGATGATGCTGCTTCAACCGTCTTCTCC 1585
Oy AAGGCGCAGAGCGGAGAAATGAAATCCCTGATGATGAGGCGCTGTGATCTTCAACCG 1080
Db AAGGCGCAGAGCGGAGAAATGAAATCCCTGATGATGAGGCGCTGTGATCTTCAACCG 1645
Oy AAGGCGCAGAGCGGAGAAATGAAATCCCTGATGATGAGGCGCTGTGATCTTCAACCG 1645
Db AAGGCGCAGAGCGGAGAAATGAAATCCCTGATGATGAGGCGCTGTGATCTTCAACCG 1705
Oy AAGGCGCAGAGCGGAGAAATGAAATCCCTGATGATGAGGCGCTGTGATCTTCAACCG 1705
Db AAGGCGCAGAGCGGAGAAATGAAATCCCTGATGATGAGGCGCTGTGATCTTCAACCG 1765
Oy CTGAGACCTGAGCTGCTCAAGTGAAGACATCCCTGAGAGGCTCTTAAACATTT 1200
Db CTGAGACCTGAGCTGCTCAAGTGAAGACATCCCTGAGAGGCTCTTAAACATTT 1765
Oy GAGGATGATCTGCTGAGCTGAGACATGATGCTTCCCTGAGAGTGTCCGACATGGTGGT 1260
Db GAGGATGATCTGCTGAGCTGAGACATGATGCTTCCCTGAGAGTGTCCGACATGGTGGT 1825

Oy GGAATTTCCGCTCTTCCAGAGAGACAGGAGCCGATGACGCTGCTGATGATGCTCTAC 1320
Db GGAATTTCCGCTCTTCCAGAGAGACAGGAGCCGATGACGCTGCTGATGATGCTCTAC 1885
Oy AAGAACCACTCTCTGAGCTTTGTGGGACCAAAAGTGGCAAGCTGAAAGATCCGGGTG 1380
Db AAGAACCACTCTCTGAGCTTTGTGGGACCAAAAGTGGCAAGCTGAAAGATCCGGGTG 1945
Oy GATGAGACCCAGGGGCAAGCCCTTCCAGATGAGACGATGAGAGCTGAGAGCTGAGAGCT 1440
Db GATGAGACCCAGGGGCAAGCCCTTCCAGATGAGACGATGAGAGCTGAGAGCTGAGAGCT 2005
Oy GTCTCTCCGGATATGAGCTTCTCCAGAGACCAAGACATCTTACATGATGAGAGAG 1500
Db GTCTCTCCGGATATGAGCTTCTCCAGAGACCAAGACATCTTACATGATGAGAGAG 2065
Oy GAGCTCACAGAGTCCCTGAGAGTCTGAGTCAATGAGAGCTGAGAGCTGAGAGCT 1560
Db GAGCTCACAGAGTCCCTGAGAGTCTGAGTCAATGAGAGCTGAGAGCTGAGAGCT 2125
Oy GAGCTCACAGAGTCCCTGAGAGTCTGAGTCAATGAGAGCTGAGAGCTGAGAGCT 2125
Db GAGCTCACAGAGTCCCTGAGAGTCTGAGTCAATGAGAGCTGAGAGCTGAGAGCT 2185
Oy GAGCTCACAGAGTCCCTGAGAGTCTGAGTCAATGAGAGCTGAGAGCTGAGAGCT 2185
Db GAGCTCACAGAGTCCCTGAGAGTCTGAGTCAATGAGAGCTGAGAGCTGAGAGCT 2245
Oy GAGCTCACAGAGTCCCTGAGAGTCTGAGTCAATGAGAGCTGAGAGCTGAGAGCT 2245
Db GAGCTCACAGAGTCCCTGAGAGTCTGAGTCAATGAGAGCTGAGAGCTGAGAGCT 2305
Oy GAGCTCACAGAGTCCCTGAGAGTCTGAGTCAATGAGAGCTGAGAGCTGAGAGCT 2305
Db GAGCTCACAGAGTCCCTGAGAGTCTGAGTCAATGAGAGCTGAGAGCTGAGAGCT 2365
Oy GAGCTCACAGAGTCCCTGAGAGTCTGAGTCAATGAGAGCTGAGAGCTGAGAGCT 2365
Db GAGCTCACAGAGTCCCTGAGAGTCTGAGTCAATGAGAGCTGAGAGCTGAGAGCT 2425
Oy GAGCTCACAGAGTCCCTGAGAGTCTGAGTCAATGAGAGCTGAGAGCTGAGAGCT 2425
Db GAGCTCACAGAGTCCCTGAGAGTCTGAGTCAATGAGAGCTGAGAGCTGAGAGCT 2485
Oy GAGCTCACAGAGTCCCTGAGAGTCTGAGTCAATGAGAGCTGAGAGCTGAGAGCT 2485
Db GAGCTCACAGAGTCCCTGAGAGTCTGAGTCAATGAGAGCTGAGAGCTGAGAGCT 2545
Oy GAGCTCACAGAGTCCCTGAGAGTCTGAGTCAATGAGAGCTGAGAGCTGAGAGCT 2545
Db GAGCTCACAGAGTCCCTGAGAGTCTGAGTCAATGAGAGCTGAGAGCTGAGAGCT 2605
Oy GAGCTCACAGAGTCCCTGAGAGTCTGAGTCAATGAGAGCTGAGAGCTGAGAGCT 2605
Db GAGCTCACAGAGTCCCTGAGAGTCTGAGTCAATGAGAGCTGAGAGCTGAGAGCT 2665
Oy GAGCTCACAGAGTCCCTGAGAGTCTGAGTCAATGAGAGCTGAGAGCTGAGAGCT 2665
Db GAGCTCACAGAGTCCCTGAGAGTCTGAGTCAATGAGAGCTGAGAGCTGAGAGCT 2725
Oy GAGCTCACAGAGTCCCTGAGAGTCTGAGTCAATGAGAGCTGAGAGCTGAGAGCT 2725
Db GAGCTCACAGAGTCCCTGAGAGTCTGAGTCAATGAGAGCTGAGAGCTGAGAGCT 2785
Oy GAGCTCACAGAGTCCCTGAGAGTCTGAGTCAATGAGAGCTGAGAGCTGAGAGCT 2785
Db GAGCTCACAGAGTCCCTGAGAGTCTGAGTCAATGAGAGCTGAGAGCTGAGAGCT 2845
Oy GAGCTCACAGAGTCCCTGAGAGTCTGAGTCAATGAGAGCTGAGAGCTGAGAGCT 2845
Db GAGCTCACAGAGTCCCTGAGAGTCTGAGTCAATGAGAGCTGAGAGCTGAGAGCT 2905
Oy GAGCTCACAGAGTCCCTGAGAGTCTGAGTCAATGAGAGCTGAGAGCTGAGAGCT 2905
Db GAGCTCACAGAGTCCCTGAGAGTCTGAGTCAATGAGAGCTGAGAGCTGAGAGCT 3460

QY 4561 AACTGTGACACCATCACTCAGGTCAAGAGAAATTCTGATGTCATCTTCAGAAATGTG 4620
 Db 5126 AACTGTGACACCATCACTCAGGTCAAGAGAAATTCTGATGTCATCTTCAGAAATGTG 5185
 QY 4621 CCTTGCTCCACCGGCCCCAAGAGCTGCAGATATGATCTGAGATGGCCCAAGAGAAATGGG 4680
 Db 5186 CCTTGCTCCACCGGCCCCAAGAGCTGCAGATATGATCTGAGATGGCCCAAGAGAAATGGG 5245
 QY 4681 GCAAGATGATCTTGCAGAGATGAAGACATCAACCAAGATTGAGATGATGGAAGCGA 4740
 Db 5246 GCAAGATGATCTTGCAGAGATGAAGACATCAACCAAGATTGAGATGATGGAAGCGA 5305
 QY 4741 CTGAACACACTGCGCCACTACCAAGGTGCCAGATGTTCCGTGTGGCATTTAGTGTCCAG 4800
 Db 5306 CTGAACACACTGCGCCACTACCAAGGTGCCAGATGTTCCGTGTGGCATTTAGTGTCCAG 5365
 QY 4801 CAGGTGACAGCCTATTAAGCAGATGAACAACCTCTCCAGAGCCTCAGCAAGTAA 4860
 Db 5366 CAGGTGACAGCCTATTAAGCAGATGAACAACCTCTCCAGAGCCTCAGCAAGTAA 5425
 QY 4861 TATGAAAACATGATCCGGTACACGGGACAGCCCTCCCTCAGCAACTATG 4920
 Db 5426 TATGAAAACATGATCCGGTACACGGGACAGCCCTCCCTCAGCAACTATG 5485
 QY 4921 ATCACTCTGACCTGAGAGATGAGATCAAGATGTGGCACTAGTGAAGAACCAAGAGAC 4980
 Db 5486 ATCACTCTGACCTGAGAGATGAGATCAAGATGTGGCACTAGTGAAGAACCAAGAGAC 5545
 QY 4981 GGAGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGATGATGTCTGAATCTACCTGACCGA 5040
 Db 5546 GGAGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGATGATGTCTGAATCTACCTGACCGA 5605
 QY 5041 CTCCTGCGCACTAAGGCGCACATGCAAGAGTTGTGATGATGATCTTTGAGAACCATCTTC 5100
 Db 5606 CTCCTGCGCACTAAGGCGCACATGCAAGAGTTGTGATGATGATCTTTGAGAACCATCTTC 5665
 QY 5101 AGCAGCGCACACCGTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5160
 Db 5666 AGCAGCGCACACCGTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5725
 QY 5161 GATGAGCAGCTGATTAACATGATGATTCATGACCCGCACTCCGCAATACCTGGAAGAGC 5220
 Db 5726 GATGAGCAGCTGATTAACATGATGATTCATGACCCGCACTCCGCAATACCTGGAAGAGC 5785
 QY 5221 AATTGCTGCGCTGAGTGGTGGTCAATGATCAAGAACCCGCACTGCTGCTGCTGCTGCTG 5280
 Db 5786 AATTGCTGCGCTGAGTGGTGGTCAATGATCAAGAACCCGCACTGCTGCTGCTGCTGCTG 5845
 QY 5281 ATCCATTAAGAACGATCAAGACGCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5340
 Db 5846 ATCCATTAAGAACGATCAAGACGCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5905
 QY 5341 TCTTGCTCAAGCTCAAGAGACCGGCTGGGCAAGGATCGCCCTCCCAAGAGCTGCTGAT 5400
 Db 5906 TCTTGCTCAAGCTCAAGAGACCGGCTGGGCAAGGATCGCCCTCCCAAGAGCTGCTGAT 5965
 QY 5401 GCCAAGAGACATCCCAAGCTCAAGAGATGGGTGGAGATTAATCAACATAGAGAGAG 5460
 Db 5966 GCCAAGAGACATCCCAAGCTCAAGAGATGGGTGGAGATTAATCAACATAGAGAGAG 6025
 QY 5461 ATCCAGGCACTCAAGAGACCAAGACCAAGACCAAGACCAAGACCAAGACCAAGACCAAG 5520
 Db 6026 ATCCAGGCACTCAAGAGACCAAGACCAAGACCAAGACCAAGACCAAGACCAAGACCAAG 6085
 QY 5521 ATGAATGATTTCAACACCAAGATGATGATCTCTGAGATCTTCTCTATGTGGGCAATAC 5580
 Db 6086 ATGAATGATTTCAACACCAAGATGATGATCTCTGAGATCTTCTCTATGTGGGCAATAC 6145
 QY 5581 AGGAGAGATCTCTTGAAGCTCTGAGACCAAGATGATGATGATGATGATGATGATGATGAT 5640
 Db 6146 AGGAGAGATCTCTTGAAGCTCTGAGACCAAGATGATGATGATGATGATGATGATGATGAT 6205

QY 5641 TACAACTGGAACAGTCAATACCTCATAGACTTGAAGAGTGA 5687
 Db 6206 TACAACTGGAACAGTCAATACCTCATAGACTTGAAGAGTGA 6252

RESULT 3
 AB073228
 LOCUS
 DEFINITION
 Mus musculus mRNA for plexin-A4, complete cds.
 VERSION
 AB073228.1 GI:28200902
 KEYWORDS
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 1
 Suto, F., Murakami, Y., Nakamura, F., Goshima, Y. and Fujisawa, H.
 Identification and characterization of a novel mouse plexin,
 plexin-A4
 Mech. Dev. 120 (3), 385-396 (2003)

JOURNAL
 MEDLINE
 22480023
 PUBMED
 12591607
 2 (bases 1 to 7022)
 Suto, F. and Fujisawa, H.
 Direct Submission
 Submitted (18-OCT-2001) Fumikazu Suto, Nagoya University Graduate
 School of Science, Chikusa-ku, Nagoya, Aichi 464-8602, Japan
 Biological Science, Chikusa-ku, Nagoya, Aichi 464-8602, Japan
 (E-mail:fumiwad@bio.nagoya-u.ac.jp, Tel:81-52-789-2981,
 Fax:81-52-789-2979)

FEATURES
 source
 Location/Qualifiers
 1..7022
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 983..6655
 /codon_start=1
 /product="plexin-A4"
 /protein_id="BAC6599.1"
 /db_xref="GI:28200903"
 /translation="MPPNMTCLLSHLVVGMSGSLLPPOPLQSGKPSFVFRGEPBA
 EGNHIVVERHTGIVLGAIVNRIYKLSLQKLVYVHTGDEEDNPCKYPIVQTECN
 PLASTNNVNMKLIDYKRNRIACSGSLGICLRLFEFLKGEFHKKEHVLGYN
 ESGSVFVIVISINFDKLFATAVDNGKREYPTTISRKLTKNSRSDMGKAAVYFHEF
 VASMIKIPDSTFVLPDPDIYYVYGSNGFVFLQLPQSVSPESTKEQVYTSKL
 VRLCKEDTAFNSVYEVIGERNVGRVRLQAAVLSKAGAVLRTGLVSPDDLTFLV
 FSKGQRKMKSLDESALCIFILKQINDRIKDRLOSQCYRGEGTDLMLKVKDIPCSA
 LRTIDNPGCLDNAAFLGVSEWVRIPIFTEDRDRTSVIAYYKNSHIAFVGTSGSK
 LKKIRVDPKGNALQYEVTVGVDSGVLDNMAKSDHEOLYIMSEKQLREVPESCG
 YRSCGCELSGSDHCGMVCYLAHNTCTKREKCEKREKREKREKREKREKREKREKREKREK
 SÖYNNLVLTIVNPELSAGVNTCTPDSMDQVYGNQICVSPAKKVPRIITENG
 DHHVQDLKSKSTGTFSTFSTFVFNCSFNSCLSVESPYCHCKKRAHNGTGRN
 TCFORQRYVLPDCCQLRVDKILVPEVETKPTLAKKLPPÖSGÖKRYBECLINQ
 KGEORVPLAFNSVQCCNTSYSYBGEINNLPELVNWNHENDINDAQKLVLY
 KGAEMESGCLCLKADPDECCGSPGQCTLQHPCHPESRMLSELSGAKNTNRI
 TEIIPVTPREGKVTIRGENLGLERDLASAVKAAVGCSEPLVDGYIPEAOINTEM
 GKAPQOHAGFVEICVAVCPREEMASQSLYEMTLLTLDKRNRGMSGGTQVITG
 TNLAGSNVYVMGSDPCLFHRSPSYIITNTSSREYITDKMTKTVQVDPRIKQIDLYE
 QYVEDPITVIEFMSIVSGNTPIAWGTHLIDLIQNPQIPAKKGEKHINI CVLNAT
 EMTQCPALALGPDHOSDLTERPEEFGFIIDNQSLILAKNTFTYIPNPEAFSPS
 GILKKEGTPILIKKRLIPVAGAVKLVYLVQEKPTVTSVDQLLCESPNLIG
 RHKMARVGEYESPGWVYIADPSPLPAIVIAVAGGILIFIAVALYARKSKLE
 SDLTILKQMONMLNBERVALBECKEAFELQTDIHETLSLDGAGIFLIDRYRYTWK
 LFGIIEHPFLRLDLYVGYQREVERGKLFQALINNKVPLSFIRITLESQSFMSD
 KGNVASTLMTVILSKLEADTVKOLLADIDKNSKXHPKLLRTEESVAEMLN
 WFTPLKPTKECQGEPLSLFCALIQOMKEKPIIDITGGAAYSLSDSLKIRQOIEKX
 TTVLSQVSPDNVSPVFPVKILNCDTITQYKELIDALIFQVNCSESRPKAADMLER
 QGSGARMILQDEDDITTKIENDMRKLNVAHYQVDPKVALVSRQVTAIVANNSTYS
 RTSASKEMNIRYTGSPDSLRSRTPMITPDESGVSKLMLVKRHEBDKQEGDGSNS
 VSEIYLRFLATGTLQKFPVDDLEFETFTAHNGSALFLAIKTMFPLDQACKGIH
 DPHVHTWKSNCPLRHWVMYIKNPQVFPDIHNSIIDACLSVAPQTFMDSGTSERH
 LKDSPEKNTLLYAKDIPSYKWTVERYSQIGKPAISDQCMNAYLAEQSRHMEFYT

ORIGIN

MSALSEIFSVYKXSEELIPLDHDQCKQKLAVKLEQVITLMSIDS"

Query Match 82.8%; Score 4714.2; DB 10; Length 7022;
Best Local Similarity 89.4%; Pred. No. 0;
Matches 5086; Conservative 0; Mismatches 588; Indels 3; Gaps 1;

QY 1 ATGAAAGCATGCGCTGGAACCTGACCTGCTCTCCACCTCCATGATGGAGATG 60
DB ATGAAAGCATGCGCTGGAACCTGACCTGCTCTCCACCTCCATGATGGAGATG 1033
QY 61 GGCCTCCACCTGCTGCTGACCGGAGCGAGCGCCGCTGCTCCAGAAAGCGCTCAT 120
DB GGCCTCCACCTGCTGCTGACCGGAGCGAGCGCCGCTGCTCCAGAAAGCGCTCAT 1090
QY 121 GTCAATTCGAGAGAGAGCGCGCGAGGTTTCATACCTGATGATGAGAGACA 180
DB GTCAATTCGAGAGAGAGCGCGCGAGGTTTCATACCTGATGATGAGAGACA 1150
QY 1091 GTGACATTCGAGAGAGAGCGCGCGAGGTTTCATACCTGATGATGAGAGACA 1150
DB GTGACATTCGAGAGAGAGCGCGCGAGGTTTCATACCTGATGATGAGAGACA 1150
QY 181 GGCACATTTACTTGGGGGCGCTCATGATTTACAGCTCTCCAGCGACCTGAGATC 240
DB GGCACATTTACTTGGGGGCGCTCATGATTTACAGCTCTCCAGCGACCTGAGATC 1210
QY 241 TTGATGACGATGAGACAGGCGCGAGAGAGACACCCAGATGTTACCCAGCGATC 300
DB TTGATGACGATGAGACAGGCGCGAGAGAGACACCCAGATGTTACCCAGCGATC 1270
QY 1211 TTGGTGAATCTCAACAGAGGCGCAGATGAGAGACACCCAGATGTTACCCAGAT 1270
DB TTGGTGAATCTCAACAGAGGCGCAGATGAGAGACACCCAGATGTTACCCAGAT 1270
QY 301 GTCCAGACCTGCAATGAGACCGCTGACACACCAACATGTCAACAGATGCTCTATA 360
DB GTCCAGACCTGCAATGAGACCGCTGACACACCAACATGTCAACAGATGCTCTATA 1330
QY 1271 GTTCAGACCTGCAATGAGACCGCTGACACCAACATGTCAACAGATGCTCTATA 1330
DB GTTCAGACCTGCAATGAGACCGCTGACACCAACATGTCAACAGATGCTCTATA 1330
QY 361 GACTCAAGAGAGACAGGCTGATGCTGTTGGAGGCTGTACCAAGCATTTGCAAGCTG 420
DB GACTCAAGAGAGACAGGCTGATGCTGTTGGAGGCTGTACCAAGCATTTGCAAGCTG 1331
QY 1331 GACTCAAGAGAGACAGGCTGATGCTGTTGGAGGCTGTACCAAGCATTTGCAAGCTG 1390
DB GACTCAAGAGAGACAGGCTGATGCTGTTGGAGGCTGTACCAAGCATTTGCAAGCTG 421
QY 421 CTGAGGCTGAGAGACCTGTTCAAGCTGGGGAGGCTTATCATAGAAAGAGACATATCTG 480
DB CTGAGGCTGAGAGACCTGTTCAAGCTGGGGAGGCTTATCATAGAAAGAGACATATCTG 1391
QY 1391 CTGAGGCTGAGAGACCTGTTCAAGCTGGGGAGGCTTATCATAGAAAGAGACATATCTG 1450
DB CTGAGGCTGAGAGACCTGTTCAAGCTGGGGAGGCTTATCATAGAAAGAGACATATCTG 481
QY 481 TCAGGTGTCAACGAGAGCGGCTCACTCTTTGAGATGATGCTCTCTCAACAGACCTGAT 540
DB TCAGGTGTCAACGAGAGCGGCTCACTCTTTGAGATGATGCTCTCTCAACAGACCTGAT 1451
QY 1451 TCAGGTGTCAACGAGAGCGGCTCACTCTTTGAGATGATGCTCTCTCAACAGACCTGAT 1510
DB TCAGGTGTCAACGAGAGCGGCTCACTCTTTGAGATGATGCTCTCTCAACAGACCTGAT 541
QY 541 GACAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
DB GACAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1511
QY 1511 GATTAACCTCTTATGACACAGAGATGATGATGATGATGATGATGATGATGATGATG 1570
DB GATTAACCTCTTATGACACAGAGATGATGATGATGATGATGATGATGATGATGATG 601
QY 601 AGCCGGAACCTGACAGAGACTGAGGCGATGATGATGATGATGATGATGATGATGATG 660
DB AGCCGGAACCTGACAGAGACTGAGGCGATGATGATGATGATGATGATGATGATGATG 1571
QY 1571 AGCCGGAACCTGACAGAGACTGAGGCGATGATGATGATGATGATGATGATGATGATG 661
QY 661 GAGTTCGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
DB GAGTTCGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1631
QY 1631 GATTTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1690
DB GATTTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 721
QY 721 GATATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
DB GATATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1691
QY 1691 GATATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1750
DB GATATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 781
QY 781 CCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
DB CCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1751
QY 1751 CCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1810
DB CCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 841
QY 841 GTGAGGCTTTGCAAGAGAGACAGCTTCACTCTTATGATGATGATGATGATGATGATG 900
DB GTGAGGCTTTGCAAGAGAGACAGCTTCACTCTTATGATGATGATGATGATGATGATG 1811
QY 1811 GTGAGGCTTTGCAAGAGAGACAGCTTCACTCTTATGATGATGATGATGATGATGATG 901
QY 901 GAGCGGATGAGGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
DB GAGCGGATGAGGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1871
QY 1871 GAGCGGATGAGGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1930
DB GAGCGGATGAGGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 961
QY 961 GTGCTTGGAGAGACCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
DB GTGCTTGGAGAGACCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG

DB 1931 GTGCTTGGAGAGACCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1990
QY 1021 AAGGCGCAG 1080
DB AAGGCGCAG 1991
QY 1991 AAGGCGCAG 2050
DB AAGGCGCAG 1081
QY 1081 AAGCAGATTAATGACCGATTAAGAGAGCGCTGACAGCTTGTATCCGAGGCGAGAGAGAG 1140
DB AAGCAGATTAATGACCGATTAAGAGAGCGCTGACAGCTTGTATCCGAGGCGAGAGAGAG 2051
QY 2051 AAGCAGATTAATGACCGATTAAGAGAGCGCTGACAGCTTGTATCCGAGGCGAGAGAGAG 2110
DB AAGCAGATTAATGACCGATTAAGAGAGCGCTGACAGCTTGTATCCGAGGCGAGAGAGAG 1141
QY 1141 CTGAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
DB CTGAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2111
QY 2111 CTGAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2170
DB CTGAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1201
QY 1201 GACGATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
DB GACGATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2171
QY 2171 GATGACAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2230
DB GATGACAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1261
QY 1261 GGAATTCCTGCTTTCAG 1320
DB GGAATTCCTGCTTTCAG 2231
QY 2231 GGCATTCAGATCTTTCAG 2290
DB GGCATTCAGATCTTTCAG 1321
QY 1321 AAGAACACTCTCTGAGCTTGTGAGGAGACAAAGATGAGAGAGAGAGAGAGAGAGAGAG 1380
DB AAGAACACTCTCTGAGCTTGTGAGGAGACAAAGATGAGAGAGAGAGAGAGAGAGAGAG 2291
QY 2291 AAGAACACTCTCTGAGCTTGTGAGGAGACAAAGATGAGAGAGAGAGAGAGAGAGAGAG 2350
DB AAGAACACTCTCTGAGCTTGTGAGGAGACAAAGATGAGAGAGAGAGAGAGAGAGAGAG 1381
QY 1381 GATGACCCAGAGGAG 1440
DB GATGACCCAGAGGAG 2351
QY 2351 GATGAGGCGCAG 2410
DB GATGAGGCGCAG 1441
QY 1441 GTCCCTCGGAGATGAGGCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
DB GTCCCTCGGAGATGAGGCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2411
QY 2411 GTCTCCGAGAGATGAGGCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2470
DB GTCTCCGAGAGATGAGGCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1501
QY 1501 CAGCTCAACAG 1560
DB CAGCTCAACAG 2471
QY 2471 CAGCTCAACAG 2530
DB CAGCTCAACAG 1561
QY 1561 GGTCTAG 1620
DB GGTCTAG 2531
QY 2531 GGTCTAG 2590
DB GGTCTAG 1621
QY 1621 CGGTGTGAG 1680
DB CGGTGTGAG 2591
QY 2591 CGGTGTGAG 2650
DB CGGTGTGAG 1681
QY 1681 CTGACGCTCCATCCCAAG 1740
DB CTGACGCTCCATCCCAAG 2651
QY 2651 CTGACGCTCCATCCCAAG 2710
DB CTGACGCTCCATCCCAAG 1741
QY 1741 ACCTACATGCTCCGAG 1800
DB ACCTACATGCTCCGAG 2711
QY 2711 ACATTAATGCTCCGAG 2770
DB ACATTAATGCTCCGAG 1801
QY 1801 ATGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860
DB ATGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2771
QY 2771 ATGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2830
DB ATGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1861
QY 1861 CCCCGATCATCAG 1920
DB CCCCGATCATCAG 2831
QY 2831 CCCCGATCATCAG 2890
DB CCCCGATCATCAG 1921
QY 1921 GAGACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1980
DB GAGACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2891
QY 2891 GAGACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2950
DB GAGACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1981
QY 1981 TCGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
DB TCGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2951
QY 2951 TCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3010
DB TCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2041
QY 2041 TGCACCAATGACCCCAAG 2100
DB TGCACCAATGACCCCAAG 3011
QY 3011 TGTACTCAGAGTGAACACATGCTCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3070
DB TGTACTCAGAGTGAACACATGCTCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

QY 2101 TGCCCCAGCTGTCGAGTGAACAAGATCTGGTGCCTGGAGGTGATCAAGCTATC 2160
Db 3071 TGCCCCAGCTGTCGAGTGAACAAGATCTGGTGCCTGGAGGTGATCAAGCTATC 3130
QY 2161 AGCTGAAGGCGAAGAACCTTCCCGAGCCCAAGTCTGGGCAAGCTGGCTACGATGATC 2220
Db 3131 ACTCTGAAGGCGAAGAACCTTCCCGAGCCCAAGTCTGGGCAAGAGGCTACGATGATC 3190
QY 2221 CTCAACATTGAGGAGAGGAGAGAGTGGCCGCTTCAACACTCTCAAGGCTA 2280
Db 3191 TTGAACATTCAGGAGATGAGAGAGGCTTCTGCGCTTCAACACTCTCAAGGCTA 3250
QY 2281 CAGTGCAGAACACCTCTTATTCCTATGAAGGAGTGAATCAACACTGCTGGAG 2340
Db 3251 CAGTGCAGAACACCTCTTATTCCTATGAAGGAGTGAATCAACACTGCTGGAG 3310
QY 2341 TTGAACATTCAGGAGATGAGAGAGGCTTCAACACTGCTGGAGTGAATCA 2400
Db 3311 TTGAACATTCAGGAGATGAGAGAGGCTTCAACACTGCTGGAGTGAATCA 3370
QY 2401 CTCTACAGTGTGAGAGGAGTGTGAGAGTGGCGCTGTCCTCAAGGCTGA 2460
Db 3371 CTCTACAGTGTGAGAGGAGTGTGAGAGTGGCGCTGTCCTCAAGGCTGA 3430
QY 2461 TTGCGATGTGCTGTGTCGAGGAGCCCAAGCTGACACCTGGGCGCAAGCTGCTG 2520
Db 3431 TTGCGATGTGCTGTGTCGAGGAGCCCAAGCTGACACCTGGGCGCAAGCTGCTG 3490
QY 2521 CAGAGAGAGCAGTGGCTGAGCTGTCTGATGCCAAAGCAAGTCAACAAACCCCGATC 2580
Db 3491 CATGAGAGCGGCTGTGAGCTGTCTGATGCCAAAGCAAGTCAACAAACCCCGATC 3550
QY 2581 ACAAGATATTCGCGGTGACAGGCCCCGAGGAGGAGGAGCAAGGCTCATATCCAGG 2640
Db 3551 ACAAGATATTCGCGGTGACAGGCCCCGAGGAGGAGGAGCAAGGCTCATATCCAGG 3610
QY 2641 GAGAAGCTGGGCTGGAATTTGCGGACATGCTGCCATGTCAGGTTGCTGGGCTGAG 2700
Db 3611 GAGAAGCTGGGCTGGAATTTGCGGACATGCTGCCATGTCAGGTTGCTGGGCTGAG 3670
QY 2701 TGCAAGCTTTAGTGAATGTTTACATCCCTGCAAGACAGATGCTGTGTGATGAGG 2760
Db 3671 TGCAAGCTTTAGTGAATGTTTACATCCCTGCAAGACAGATGCTGTGTGATGAGG 3730
QY 2761 GCGAAGCCCAAGCAGATGAGAGGCTTGGTGAAGTGGCGGTGTGTGTGAGGCTGAG 2820
Db 3731 GCGAAGCCCAAGCAGATGAGAGGCTTGGTGAAGTGGCGGTGTGTGTGAGGCTGAG 3790
QY 2821 TTCAATGCGCGGTCTCAAGCTCTATTACTTCAATGACACTGACTCTCTCAATCTGAG 2880
Db 3791 TTCAATGCGCGGTCTCTCAAGCTCTATTACTTCAATGACACTGACTCTCTCAATCTGAG 3850
QY 2881 CCCAGCGGAGGAGCCCAATGCTCGAGAGGAGCCCAATGACATCAAGGAGCCCAATGAT 2940
Db 3851 CCCAGCGGAGGAGCCCAATGCTCGAGAGGAGCCCAATGACATCAAGGAGCCCAATGAT 3910
QY 2941 GCGGAGAGCAAGTGTGTGTGTGATGTTTGAAGAGAGCCCTGCTCTTCAAGGAGTCT 3000
Db 3911 GCGGAGAGCAAGTGTGTGTGTGATGTTTGAAGAGAGCCCTGCTCTTCAAGGAGTCT 3970
QY 3001 CCATCTTCAATTTGTGCAACCAATCTCTCAAGTGTGTGTGTGTGTGTGTGTGTGTGT 3060
Db 3971 CCATCTTCAATTTGTGCAACCAATCTCTCAAGTGTGTGTGTGTGTGTGTGTGTGTGT 4030
QY 3061 GTGAGGTGAGCAGGAGCCCAATGCAAGGAGCCCTGCTCTTCAAGTGTGTGTGTGTGTGT 3120
Db 4031 GTGAGGTGAGCAGGAGCCCAATGCAAGGAGCCCTGCTCTTCAAGTGTGTGTGTGTGTGT 4090
QY 3121 ACCATCTGTGCGGATTTGAGCCAGATGAGAGCATTTGTCACTGTGAACACACCCTATGCGGTA 3180
Db 4091 ACCATCTGTGCGGATTTGAGCCAGATGAGAGCATTTGTCACTGTGAACACACCCTATGCGTCTC 4150

QY 3181 TGGGGAGCCCACTTGAGCTCATACAGAACCCCGAGATCCGTGCTCAAGCATGAGGAG 3240
Db 4151 TGGGGAGCCCACTTGAGCTCATACAGAACCCCGAGATCCGTGCTCAAGCATGAGGAG 4210
QY 3241 GAGCATCAATATTTGAGAGGTTTGAAGCTATGAGATGAGACCTGTGAGGCGCGCC 3300
Db 4211 GAGCATCAATATTTGAGAGGTTTGAAGCTATGAGATGAGACCTGTGAGGCGCGCC 4270
QY 3301 CTGCTCTGGGTCTGACCAAGTCAAGCTGACCGAGAGGCGCGAGAGTTTGGCTTC 3360
Db 4271 CTGCTCTGGGTCTGACCAAGTCAAGCTGACCGAGAGGCGCGAGAGTTTGGCTTC 4330
QY 3361 ATCTGAGCAGATGCTGCT 3420
Db 4331 ATCTGAGCAGATGCTGCT 4390
QY 3421 AACCGGTGTTGAGGCTTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3480
Db 4391 AACCGGTGTTGAGGCTTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4450
QY 3481 ATCATCTTAAGGAGCAAGCTGATCCGCTGTGCTGGGAGCAAGTGAAGCTGAC 3540
Db 4451 ATCATCTTAAGGAGCAAGCTGATCCGCTGTGCTGGGAGCAAGTGAAGCTGAC 4510
QY 3541 TACATGTGCTGTGTTGGGAGAGAGCGTGAACCGTGAACCGTGTGATGTCCAGTGTCTC 3600
Db 4511 TACATGTGCTGTGTTGGGAGAGAGCGTGAACCGTGAACCGTGTGATGTCCAGTGTCTC 4570
QY 3601 TGCAGTCCCACTCAATGCGAGGAGCAAAAGTATGCGCTGTGCTGTGCTGTGCTGTG 3660
Db 4571 TGCAGTCCCACTCAATGCGAGGAGCAAAAGTATGCGCTGTGCTGTGCTGTGCTGTG 4630
QY 3661 TACTCCCGGAGATGTGTACATGCCCCGAGCAAGCTGAGCTGAGCTGAGCTGAGCTG 3720
Db 4631 TACTCCCGGAGATGTGTACATGCCCCGAGCAAGCTGAGCTGAGCTGAGCTGAGCTG 4690
QY 3721 AGCATGAGTGGCTGTGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3780
Db 4691 AGCATGAGTGGCTGTGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4750
QY 3781 AAACGCAATCCCGGAGATGTGTACATGCCCCGAGCAAGCTGAGCTGAGCTGAGCTG 3840
Db 4751 AAACGCAATCCCGGAGATGTGTACATGCCCCGAGCAAGCTGAGCTGAGCTGAGCTG 4810
QY 3841 CTGAGTCCCGTGTGAGGCTGTGAGAGGAGCTTGTGAGCTGAGCTGAGCTGAGCTG 3900
Db 4811 CTGAGTCCCGTGTGAGGCTGTGAGAGGAGCTTGTGAGCTGAGCTGAGCTGAGCTG 4870
QY 3901 CATGAGCTACCAATGAGCTGAGTGAAGCGGAGATTCGTTCTGTGATGATGAGCTTAC 3960
Db 4871 CATGAGCTACCAATGAGCTGAGTGAAGCGGAGATTCGTTCTGTGATGATGAGCTTAC 4930
QY 3961 ACCATGCGGATGTCT 4020
Db 4931 ACCATGCGGATGTCT 4990
QY 4021 CCGAGGTACCGAGAGGCTGTGAGAGAGGCTGAGGCTTGTGAGCTGAGCTGAGCTG 4080
Db 4991 CCGAGGTACCGAGAGGCTGTGAGAGAGGCTGAGGCTTGTGAGCTGAGCTGAGCTG 5050
QY 4081 AACAGGTGTCTGTGCT 4140
Db 5051 AACAGGTGTCTGTGCT 5110
QY 4141 CCGAGCTGTGAGAGGCTGTGAGAGGCTGTGAGAGGCTGTGAGAGGCTGTGAGAGG 4200
Db 5111 CCGAGCTGTGAGAGGCTGTGAGAGGCTGTGAGAGGCTGTGAGAGGCTGTGAGAGG 5170
QY 4201 GCGACTGATGTGTGAGAGGCTGTGAGAGGCTGTGAGAGGCTGTGAGAGGCTGTGAGAG 4260
Db 5171 GCGACTGATGTGTGAGAGGCTGTGAGAGGCTGTGAGAGGCTGTGAGAGGCTGTGAGAG 5230
QY 4261 AACCACTTAAGCTGTCT 4320

ORIGIN

Query Match 51.7%; Score 2939.6; DB 6; Length 6387;
Best Local Similarity 71.3%; Pred. No. 0;
Matches 3952; Conservative 0; Mismatches 1579; Indels 15; Gaps 5;

151 TTCAATCAGCTGTGTGATGAGAGACAGACATTTACTTGGGGGCGCTCAATCG 210
421 TTCAACACTTGAACCGTCCACCAAGGAGCGGGGCGGTCTATGTGGGGGCAATCAACGG 480
211 ATTTCAGAGCTTCCAGCGACCTGAAGCTTGTGTACGATGAGACAGGCGCGAGCG 270
481 GTCTATAGCTGACAGGCAACCTGACCATCAGGTGGCTCTATAAGACAGGGCCAGAAAG 540
271 GACACCCCAAGTGTATCCCAACCGCATGTGCAAGCTGCAATAGCCCTGACACAC 330
541 GACACAAAGCTTGTATCCCGCCCTCATGTGTGACGCTGACGGAAGTCTCAACCTC 600
331 ACCAACAATGTCAACAAGATGTCTCTATAGACTTACAGAGAAACAGGCTGATGCTGT 390
601 ACCAACAATGTCAACAAGATGTCTCTATAGACTTACAGAGAAACAGGCTGATGCTGT 660
391 GGGAGCGCTGTACCAAGGCACTGTGCAAGCTGTGAGGTGAGGAGCTCTTCAAGCTGG 450
661 GGGAGCTTACCAAGGCGGCTGTGCAAGCTGTGAGGTGAGTACCTTCTTCACTGAGTG 720
451 GAGCCTTATATAGAAGAGAGACATATCTGTCAAGGTGTCAACGAGAGCGGCTCACTTT 510
721 GAGCCTTATATAGAAGAGAGACATATCTGTCAAGGTGTCAACGAGAGCGGCTCACTTT 780
511 GAGATATGTCTTCCATACAGAACCTGTGAGCAAGCTGTTCATCCACGAGAGAT 570
781 GGGGTATGTGTGCGCTGTAGGAGTGAAGATGTCTTCACTGAGAGCTGTGAGT 840
571 GGGAGCGCGAGATATTTCCCAACATCTCCAGCGGAGAACTGACCAAGAACTGTGAGG 630
841 GGGAGGAGATTAATCTCCCGACCTGTGTCCAGCGGAGAGTGTCCCGAGACCTGTGAGT 900
631 GATGGCATGTTCGGATACGCTTCCATGATGATGTGTGAGCTGTGATGATTAAGATCCT 690
901 TCAGCCATGTGTCACTATGAGCTACACAGGATTTGTCTCTCTCAACAGATCCCT 960
691 TCAGCACTTCAACATCATCCCTGACTTGTATATCTACTATGTATGATGATGATGAT 750
961 TCAGCACTTCAACATCATCCCTGACTTGTATATCTACTATGTATGATGATGATGAT 1020
751 GGGCACTTGTCTACTTTTGAACCTTCAACCTGAGAA--TGGTGTCTCAACAGGCTTC 807
1021 GGGGCGTGTGTCTACTTTCTCACTGTCCAGCCGAGAACCTGTGAGGATGTGCAATCAAC 1080
808 ACCACCAAGAGAGAGGTGTATACATCAAGCTGTGAGGCTTGTGCAAGAGAGACAGGC 867
1081 TCCGCTGAGAGACCTCTTCTACACTCAAGCATGTGTGAGGCTGTGCAAGAGATGACCCCAAG 1140
868 TTCAACTCTATGTAGAGGTGCTCATGTGAGCGAGTGTGAGGATGATGATGATGATGAT 927
1141 TTCAACTCTATGTAGAGGTGCTCATGTGAGCGAGTGTGAGGATGATGATGATGATGAT 1200
928 CTGAGAGCTGAGTACCTGTGCAAGAGGCGGAGGCTGTGAGAGAGAGAGAGAGAGAGAG 987
1201 CTGAGAGCTGAGTACCTGTGCAAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260

LIILIVILIAVKRSRENDLTLKRLQOMONLBSRVALCEKEAFELQDINELTS
DLRSGPIYDIYTYTAMRVLPFGIEDHPVIRELEVANGQOYERAKLPAQILINKKY
FLILFRTLELORSFSWRDGNVASLIMGLGRLVADVLQILSDILDKNLEKN
HPLILRLTESVAEKLTWPAFLHFLHLEAGBPLFMYCAIKQOMKGPIDALTG
EAYYSISEDLIRQOIEYKTLINCVNPDENBEPJPKVLNDDITVQKEHILAVY
KNVYSORPAVMDLEMRQARVAVLOEDITTKLEGKMLNTLMHYQSDSVV
ALVFKOTSNIYPASAGISRTSISRYSRSTYSGPDSISRAPMTTPDLBSGVKWH
LVKNDHGDQKEDRGSKWSEYIYRLATKGLTOKPVDLPELTFTSTVHGSALPL
AIKMPEDLEQADBRSHIDTVRHTKSKCLPBRVANSYKPOVPIHSGSITDA
CLSVAGTENDSCSTSEHRLKDSPSKLIYADIPSYVSWERYIADIAKIPALSDQ
DMNAVLAEGRLHAVERFNMALNEIYSVSKSEBLIGALBDEQARRORLAYKEQ
LINAMSTIES"

988 CCAGATGATACCTGCTCTTCAACCGTCTTCTTCACAGGCGCAAGAGCGGAATAATGCC 1047
1261 AGCCAGAGAGATGATGATCTTTTGGCACTCTTCCCAAGGGCGAGAGAGATACACACCG 1320
1048 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1107
1321 CCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
1108 CGGCTGAGTCTTGTTCACGAGGCGAGAGGACAGCTGTGAGCTGTGAGCTGTGAGGAT 1167
1381 CGCTGAGTCTTGTTCACGAGGCGAGAGGACAGCTGTGAGCTGTGAGCTGTGAGGAT 1440
1168 GACATCCCTGAGAGAGGCTCTTACATGATGATGATGATGATGATGATGATGATGAT 1227
1441 GAGCTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
1501 AACGAGCGCTGTGAGGCTCAACTCAAGTGAAGGCGCTGACCTGTACACACAGAGAG 1560
1288 GACGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1347
1561 GACGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
1348 ACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1407
1621 ACTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
1408 TATGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1464
1681 TACGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
1465 AAGGACCAAG 1524
1741 ATTGATGAG 1800
1525 TCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1584
1801 TCATGTGAG 1860
1585 TGTGTGTGAG 1644
1861 TGTGTGTGAG 1920
1645 GCGAGGTTGCTGTGAG 1704
1921 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
1705 TCGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1764
1981 TCAGTATGAG 2040
1765 GCTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1824
2041 GGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
1825 CAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1884
2101 CAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2157
1885 GACCACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1944
2158 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2217
1945 ACCAGCTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2004
2218 ACCAGCTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2277
2005 CCATACGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2064
2278 GCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2337

QY 2065 TCCCTCCAGAAAGCCGAGTGAAGCTGCCGAGAGCTGCCCACTGCTGCGAGTGCAC 2124
DB 2338 TCCCTCCAGAAAGCCGAGTGAATTTTCAAGAGCTGCCCACTGCTGCGAGTGCAC 2397
QY 2125 AAGATCTCGTCCCGTGAAGTGAATCAAGCTTACAGCTGAAGGCCAAGACCTGCC 2164
DB 2398 GAGATCTTATTCAGATCGGAGAGTAAAGCCATACCTTTAAGCGCGAATGTGCC 2457
QY 2185 CAGCCCACTCTGAGCAGCTGAGTCAAGATGCACTTCAACATTCAGAGCGAGCAG 2244
DB 2458 CAGCCCACTGAGCGAGCGAGGCTATGATGTGCTTCAACATTCAGAGCGAGCAG 2517
QY 2245 CAGATGCCCGCTGCTTCAACAGCTTCAAGCTGCAAGTGCACAAACCTTTATTC 2304
DB 2518 CAGATGCCCGCTGCTTCAACAGCTTCAAGCTGCAAGTGCACAAACCTTTATTC 2577
QY 2305 TATGAAGGATGAGATCAACAACCTGCCGAGTGAACAGTGTGTGAATGGGAC 2364
DB 2578 TATGATGAGATGAGATCAACAACCTGCCGAGTGAACAGTGTGTGAATGGGAC 2637
QY 2365 TTCAACATTTGAACACCCAGCTCAAGATTAAGTTTCACTTCAAGTGTGAAGCCA 2424
DB 2638 TTCAATATTGAACACCCAGCTCAAGATTAAGTTTCACTTCAAGTGTGAAGCCA 2697
QY 2425 GAGAGCTGCGGCTGCTGCTCAAGGCTGACCCAGCTTGCATGTGCTGCGAGGAC 2484
DB 2698 GAGAGCTGCGGCTGCTGCTCAAGGCTGACCCAGCTTGCATGTGCTGCGAGGAC 2757
QY 2485 CAGAGCTGAGACCTGCGGAGCACTGCTGCCAGAGAGCCAGTGTGAGAGTGC 2544
DB 2758 GAGAGCTGAGACCTGCGGAGCACTGCTGCCAGAGAGCCAGTGTGAGAGTGC 2817
QY 2545 TCTGAGGCAAAAGCAAGTGCACAAACCCCGCATCAAGAGATATCCCGTGAACAG 2604
DB 2818 TCCAGCAACATGTCAAGTGTCCAACTTCAATTCAGAGATTTTGAAGTGTCTGA 2877
QY 2605 CCCCCGAAAGGGGAGCCAAAGTCACTATCCAGAGGAGAACTGAGGCTTGAATTC 2664
DB 2878 CCGCCGAGAGAGGGAGCGAGATGACATCCATGCGGTGAACCTGGGCTGGAATTC 2937
QY 2665 GAGATGCTCCCATGTCAGAGTGTGCTGCGAGTGTGAGAGGCTTGAAGTGTGAC 2724
DB 2938 GAGATGCTCCCATGTCAGAGTGTGCTGCGAGTGTGAGAGGCTTGAAGTGTGAC 2997
QY 2725 ATCCCTGCAAGACAGATCTGTGTGATGATGAGAGGAGGAGCCAGCCAGCATG 2784
DB 2998 ATCATGCTGAGAGATGTGTGTGATGAGAGGAGGAGCCAGCCAGCATG 3057
QY 2785 TTGTGTGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2844
DB 3058 CCAATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3117
QY 2845 TATTACTTATGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2904
DB 3118 TACACTTGTGATACCTTCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3177
QY 2905 GAGAGCCAAAGTACATCAAGGACCAACCTGAAATGCGGAGAGCAAGTGTGTGAT 2964
DB 3178 GAGAGCCAAAGTACATCAAGGACCAACCTGAAATGCGGAGAGCAAGTGTGTGAT 3237
QY 2965 TTGAGAAAGAGCCCTGCTGCTTCCAGAGGATCTCATCTTCAATGTGTG---CAAC 3021
DB 3238 CTGGGAGACCAAGCTGCGAGTTCACGAGAGTCAAGAGTGAATGTGTGTCTCA 3297
QY 3022 ACCACATCTCTGAGATGAGTGTCTAGATGAAGTGTGTGTGTGTGTGTGTGTGTGT 3081
DB 3298 CCCCCATCAATCAATGAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3357
QY 3082 ATCCACCAAGAGCTGCTGCTTTCAGTATGAGAAAGCCCAACCATGTGTGTGTGTGT 3141
DB 3358 GTGGATAGCAACTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3417
QY 3142 GAATGAGCATTTGTCTAGTGAACACACCCATGCGCTGTATGAGGAGACCACTGAG 3201

DB 3418 GAGTGAAGCATTTCCAGATGAGGACACACACCTTCAACATGAGCTTCAACCTGAGTGC 3477
QY 3202 ATTCAGAAACCCCAAGATCCGCTGCAAGCATGAGGAGAGAGACATCATATCTGTGAG 3261
DB 3478 ATTCAAGAGCCAAAGATCCAGTCAATTCATTAAGGCAAAAGATCTGTGTGTGTGT 3537
QY 3262 GTTCTGAACGCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3321
DB 3538 GTTCTGAACGCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3597
QY 3322 CAGTCAAGCTGACCGAGAGGCGCGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3381
DB 3598 GAGCTGAGACCTGTGAGACCGCAGATGATGATGATGATGATGATGATGATGATGATG 3657
QY 3382 CTGTCTATCTTCAACAAAGCAACCTTCACTTATCTTCAACCCGAGTGTGTGTGTGT 3441
DB 3658 TTGCTATTTTCAACAAAGCAACCTTCACTTATCTTCAACCCGAGTGTGTGTGTGT 3717
QY 3442 GGTCCCTCAGAGATCCCTGAGCTCAAGCTGAGAGCCGCTGAGATCTTAAAGGCAAGAC 3501
DB 3718 AGCTCTATGAGATCTTGTATCAAAAGCAGAGATGCTCCATCATCTTAAAGGCAAG 3777
QY 3502 CTGATCCGCTGT 3561
DB 3778 CTGTGCTCTCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3834
QY 3562 AAGCTGTGACCGTGAACCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 3621
DB 3835 ACCCTGT 3894
QY 3622 GGCAGGCAAAAGTGAATGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3681
DB 3895 GGCAGGCAAAAGTGAATGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3954
QY 3682 ATTGCCCCGAGACCGCTGACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3741
DB 3955 GTCATCTCAACAGCTTGT 4014
QY 3742 CTCTCTCATATTTTCAATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3801
DB 4015 CTCTCTCATATTTTCAATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4074
QY 3802 GACCTCAAGCTGAGAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 3861
DB 4075 GACCTCAAGCTGAGAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 4134
QY 3862 GAGTGAAGGAGAGCTTGTGCGAGCTGAGAGGAGCATCATGAGCTGAGAGGAGCTG 3921
DB 4135 GAGTGAAGGAGAGCTTGTGCGAGCTGAGAGGAGCATCATGAGCTGAGAGGAGCTG 4194
QY 3922 GATGAGAGGAGAGCTTGTGCGAGCTGAGAGGAGCATCATGAGCTGAGAGGAGCTG 3981
DB 4195 GACCGCTCAGAGATCTTGTGCGAGCTGAGAGGAGCATCATGAGCTGAGAGGAGCTG 4254
QY 3982 GGAATGAAGAGAGAGCTTGTGCGAGCTGAGAGGAGCATCATGAGCTGAGAGGAGCTG 4041
DB 4255 GGAATGAAGAGAGAGCTTGTGCGAGCTGAGAGGAGCATCATGAGCTGAGAGGAGCTG 4314
QY 4042 GTGAGAAAGAGCTTGTGCGAGCTGAGAGGAGCATCATGAGCTGAGAGGAGCTG 4101
DB 4315 GTGAGAAAGAGCTTGTGCGAGCTGAGAGGAGCATCATGAGCTGAGAGGAGCTG 4374
QY 4102 TTGATCGGAGAGCTTGTGCGAGCTGAGAGGAGCATCATGAGCTGAGAGGAGCTG 4161
DB 4375 TTGATCGGAGAGCTTGTGCGAGCTGAGAGGAGCATCATGAGCTGAGAGGAGCTG 4434
QY 4162 TCACTCATATGACCTGTGTGAGAGCAACCTGAGATGAGCACTGATGTGTGAGAGAG 4221
DB 4435 TCGCTCATATGACCTGTGTGAGAGCAACCTGAGATGAGCACTGATGTGTGAGAGAG 4494
QY 4422 CTGTGCGAGAGCTTGTGAGAGCAACCTGAGATGAGCACTGATGTGTGAGAGAG 4281

Db 4495 CTGCTCTCTGACCTCATCGATTAAGAACCTGGAGAACCAACCCCAAGCTGCTATC 4554
 QY 4282 AGGAGACGTAGTCACTGAGTGTGAGAAAGATGTCACCAATGGTTTACTTTCCTCTCTAC 4341
 Db 4555 CGAGAGACAGAGTGTGTGGCTGAAAAAGATGTCACCAATGGTTTCTCTCTCTGAC 4614
 QY 4342 AAGTTCTCAAGAGTGTGTGGAGCCCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4401
 Db 4515 AAGTTCTCAAGAGTGTGTGGAGCCCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4674
 QY 4402 CAGATGAGAGAGGAGCCCATTTGACGCCATCAAGGAGGAGCCGCTACTCTCTGAGCGAG 4461
 Db 4675 CAGATGAGAGAGGAGCCCATTTGATGCCATCAAGGAGGAGCCGCTACTCTCTGAGCGAG 4734
 QY 4462 GACAACTCATCGGCGACAGAGATTGACTAACAACCTGCTCTGAGCTGTGTACGCCA 4521
 Db 4735 GACAACTCATCGGCGACAGAGATTGACTAACAACCTGCTCTGAGCTGTGTACGCCA 4794
 QY 4522 GACAACTCATCGGCGAGGAGTGTGCTGTAAGATGCTCAACTGTGACACATCACTACG 4581
 Db 4795 GACAACTCATCGGCGAGTGTGCTGTAAGATGCTCAACTGTGACACATCACTACG 4854
 QY 4582 GTCAAGAGAGAGATTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 4641
 Db 4855 GTCAAGAGAGAGATTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 4914
 QY 4642 GCTCAAGATGATGATCTGAGTGTGCTGACAGAGAGAGAGAGAGAGATCTTTCGAGAT 4701
 Db 4915 GCAATGACATGATGCTGAGTGTGCTGACAGAGAGAGAGAGAGAGATCTTTCGAGAT 4974
 QY 4702 GAAGACATCAACCAAGATTGAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4761
 Db 4975 GAGAGACATCAACCAAGATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5034
 QY 4762 CAGGTGCCAAGTGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4821
 Db 5035 CAGGTGTCAAGACAGAGT 5094
 QY 4822 GTGACAACTCCACGCTCTCAGAGACTGACAGAGATGATGATGATGATGATGATGATG 4881
 Db 5095 CTTGCTCTGCGAGCATCTCCCGAGCTCATCAGAGAGATGATGATGATGATGATGATG 5154
 QY 4882 ACCGGGAGAGCCCGACAGCTTCTGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4941
 Db 5155 ACCGGGAGAGCCCGACAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5214
 QY 4942 GAGTCAAGAGATGAG 5001
 Db 5215 GGGGTCAAGAGATGAG 5274
 QY 5002 CGGGGAGAGAGATGAG 5061
 Db 5275 CGGGGAGAGAGATGAG 5334
 QY 5062 CTGACAGAGATGAG 5121
 Db 5335 CTGACAGAGATGAG 5394
 QY 5122 GCCCTGAG 5181
 Db 5395 GCTCTGAG 5454
 QY 5182 GGCATTCAAG 5241
 Db 5455 AGCATCATGACAG 5514
 QY 5242 TGGGTCAAGATGATCAAG 5301
 Db 5515 TGGGTCAAGATGATCAAG 5574
 QY 5302 GAGCCTGAG 5361
 Db 5575 GAGCCTGAG 5634

QY 5362 CGGCTGGGCAAGAGATCGCCCTCCCAACAGCTGTGTATGTCAGAGAGATCCGAGTAC 5421
 Db 5635 CGGCTGGGCAAGAGATCGCCCTCCCAACAGCTGTGTATGTCAGAGAGATCCGAGTAC 5694
 QY 5422 AAGATTTGGGTGAGAGATTTATCTCAGACATAGAGAGAGAGAGAGAGAGAGAGAGAG 5481
 Db 5695 AAGATTTGGGTGAGAGATTTATCTCAGACATAGAGAGAGAGAGAGAGAGAGAGAGAG 5754
 QY 5482 GACATGAACGATACCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5541
 Db 5755 GACATGAACGATACCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5814
 QY 5542 AGTGCATCTCAGAGATCTTCTCTATGTGGGCAATACAGAGAGAGATCTTGTGACT 5601
 Db 5815 AGTGCATCTCAGAGATCTTCTCTATGTGGGCAATACAGAGAGAGATCTTGTGACT 5874
 QY 5602 CTGACACAGATGACAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5661
 Db 5875 CTGACACAGATGACAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5934
 QY 5662 ACCCTGATGAGCTTACAG 5687
 Db 5935 AATGCCATGCTCATGAG 5960

RESULT 5
 AB007932
 LOCUS Homo sapiens mRNA for KIAA0463 protein, partial cds.
 DEFINITION
 AB007932.1 GI:3413887
 VERSION
 KIAA0463 protein.
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 9458484
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 FEATURES
 source
 1..6263
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="1"
 /clone="HG0942"
 /sex="male"
 /tissue_type="brain"
 /clone_id="pBluescriptII SK plus"
 1..6263
 /gene="KIAA0463"
 <1..5894
 /gene="KIAA0463"
 /codon_start=3
 /product="KIAA0463 protein"
 /protein_id="BAA3308.1"
 /db_xref="GI:3413887"
 /translation="GFGRLPDSRLRAGRSRRPQQAARVDAAGTGGTDAVLES
 CEGQKAVQMSHRSRLTLPAISMQRKRPALVDSDSVYLIVAKTLPAPRA
 GMPQSTFHSNRMTEFNHLLTVHOGTAIVVGAALNRYKLIGNLITIOVAHKTGPEEDN

KACVPLIYQPCSEVLTITNNKLLIDYSENRLIAGCSLYQVCKLRIDPLIIV
 EBSHKEHLYSSVNTKGTWYIVASBEGDKLFTIGIADGKODIFPLLSERKPRDP
 ESSAMLDYELSHDVFSLIKRIPSDTLAVSHPDIFYITGFASSGVLVYLPVOPFES
 VAINSAGDLFYTSRIVRLCKDPEKHSVYSLFPFGCTGAGVRLQAAVYLAKEPSSLA
 QAFNITSODVLPALFSPKQKQYHHPDDPSALCAPIAIAINQIGRIQSCYQSGNL
 OANLWIKDVOCCTKAPVPIIDNFCGLIDNQPGLSPVEGLTYTSSDRMTSVASYV
 YNGSYVFPVGTSGKRLKIRADGPBGVGYEMVSLDGLIDDMAFSIDORIYV
 MSEROYTRVPSKSCROYTCCGCTCSGDPICGMCALHMCSCRDCCOAMBNPRAAS
 ISBOCVLAHPSSISISESRILSLVSNAPPLSAGIACBENLREYSGVSGVTC
 ISGEPDVPIPLDQWFELEJQKSKETGKIFSTERFENKASHQCLCSNARE
 CHWCKRYNLCTHDPITCSFQGRINISEDCPOLVPEBELLFVGAVKQITLAEARLPQ
 POSGQREYECVNIQGAIRHVPALRENSSVQCONSVOYDMDLSINAVPAVWNG
 NFIIIDNPOLKYLKCAQRESCELCADRFECGMCGERCTDHOCHSPSPM
 LDMSSHNSQNSPOITELTVSGPEGGRTVTHGNLGLDSETHAHVOYAGVCTP
 LBPEVTAIOVCEMGHALVGTSGEPVRLCIGCKDEHMTKQOQYVPSVSLNP
 TRGPESGMTVITTHGYILAGSSVANYLIDNCEPFGMSITVCSPPSSNGSPVP
 VSVSVRAHVDNMLQPEYIDDPVOKIIBENKSIAGHTPLITGNRLVQDPRRVK
 FNKESVNVCKVNTTTLTCLAPSLTIDRPGIDVERDEGFVFNNOVOLLINDY
 KEIYENPTEFELSPGVLDOKRGPDIILKGNLCPBPAGGAKMYVLIIGETPAVT
 VSETOLCPEPNLTGQKYMVHVGVFPGSVSYISDLSLTPAIVISAAGSLLI
 IVTILVAYRKRENDLTLELOMOMDLSEKVALCEKEAPAEIOTDINELTSLDR
 SCIPLYDVTYAMRULPGIIEHPULREIYVQNGOHEKALKEIQAOLINNKVILT
 FIRTLEQSFMSKRDGNVASYLMTGLORLBYADYKQILSDIIDRLKEXHPKY
 LIRTESVAEKMLTNFAPULHFKPEKCEBEPFMYLCATKQOMEKPIIDATVGNAY
 YSOPRAVIMDLERQRIARVVLQDEDITPKIEGKMLNTIMHYQVDSRVVALVP
 KOTSSYNIIPASISIRTSISRYDSFRYGSPLSLAPMPTDLESVGRKWLHVN
 HDHDOKEGRSGKMSYSEIYELTAKTLOKFDVDELETFSVTHRSALPLIKY
 MPDPLDEQADRSIHDTDVIRHTKSNCLPLRFMYNTIKPOFIDPIKHSITDACLAV
 VAOTPDSCSTSEHRLGKXSPSKLIXADIDISYKSVRYRYADIAXIPLAISDOMMA
 YLAEQRLHAHVEFNMLSLALNELISYVSKISELIGLEDOEQRORLAIVKEOLINA
 MSIES"

ORIGIN

Query Match: 51.6%; Score 2938; DB 9; Length 6263;

Best Local Similarity: 71.2%; Pred. No. 0;

Matches 3951; Conservative 0; Mismatches 1580; Indels 15; Gaps 5;

151 TTCAATCAGCTGTGTGATGAGAGACAGACATTTCCTGGGGCCCTCATCGG 210
 Db TTCAACACCTTACCGTCCACCAAGGAGCGGGCCCTCTATGTGGGGCCATCAACCGG 416
 357 211 ATTACAGCTCCCGACGACCTGAAGCTGTGTAGCGATGAGACAGGGCCGAGCGAG 270
 Db 417 GTCTATAGCTGAGAGGAACTGACCATCAGGTGCTATAGACAGGGCCAGAGAG 476
 Qy 271 GACAAACCCCAAGTGTACCCACCCCGCATCGTCCAGACCTGCAATGAGCCCTGACACC 330
 Db 477 GACAAACAGTCTGTACCCGCCCTCATCGTGAAGCCCTGACAGCAAGTGTCAACCCCTC 536
 Qy 331 ACCAACAATGTCAACAAGATGCTCCATAGACTACAGAGAAAGAGGAGGCTGATGCTGT 390
 Db 537 ACCAACAATGTCAACAAGGCTGTCTATGACTACTCTAGAAACCCCTGCTGGCCCTGT 586
 Qy 391 GGGAGCCCTGTACCAAGGATCTGCAGCTGTGAGGCTGAGAGACCTCTTCAAGCTGGG 450
 Db 597 GGGAGCCCTGTACCAAGGGGCTGTGAGCTGTGAGGCTGAGAGACCTCTTCAAGCTGGG 656
 Qy 451 GAGCTTATCATTAAGAGAGACATATCTGTCAAGTGTCAACGAGAGGCGCTCAAGCTT 510
 Db 657 GAGCATCCCAAGAGAGAGACATCTGTCCAGTGTCAACAGAGCGGACCAACATGTAC 716
 Qy 511 GAGAGTATCTGTCTCTACAGCAACCTGTGATGACAGAGTGTTCATTGCGACGAGTGGAT 570
 Db 717 GGGGTGATGTGCGCTGTGAGGGTGAAGATGCAAGCTTCCTCATCGGCAAGCTGTGAT 776
 Qy 571 GGGAGAGCCGAGTATTTTCCCAACATCTCCAGCCGAAACTGACCAAGAACTCTGAGCG 630
 Db 777 GGGAGAGAGATTAATCTCCACACCTGTCCAGCCGAAAGCTGCCCGAGACCTGATCC 836
 Qy 631 GATGCGATGTTCGCGTACGCTCTTCATGATGAGTTCGTGAGCTGATGATTAAGATCCCT 690
 Db 837 TCAGGCACTGTGACTATAGACTACAGGATTTTGTCTCTCTCTCACTAAGATCCCT 896

Qy 691 TCAGACACCTTCAACACATACCTCCGACTTGTATATCTACTATGTTTGTAGAGT 750
 Db TCAGACACCTTGGCCCTGTGCTCCACTTTGACATCTTCACTACATCAGGCTTGTGAT 956
 Qy 751 GGCACATTTGTACTTTTGTGACCCCTCAACCTGAGA--TGTGTCTCCACAGGCTCC 807
 Db 957 GGGGCTTTGTACTTTTCACTGTCCAGCCGAGACCCCTAGAGGCTGTGGCCATCAAC 1016
 Qy 808 ACCACCAAGAGACAGTGTATATCATCCAACTGTGTAGAGGCTTGAAGAGACACAGCC 867
 Db 1017 TCCGTGAGACCTCTTCACTCACTCAAGCATGTGTGGCTTCTGACAGATGACCCCAAG 1076
 Qy 868 TTCAACTCTATATGAGAGTGCCTATGCTGTGAGCGCAGTGGGTGAGTACCGCTG 927
 Db 1077 TTCAACTATACGTGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1136
 Qy 928 CTGACAGCTGCTTACCTGTCCAAAGCGGGGGCGGTGCTTGGACGACCTTGGAGTCCAT 987
 Db 1137 CTGACAGCTGCTTACCTGTCCAAAGCGGGGACTCACTGACCCAGGCTTCAATATAC 1196
 Qy 988 CCAGATGATGACCTGCTCTTCAACCGCTTCTCCAGGGCCAGAACCGAAATGAAATCC 1047
 Db 1137 AGCCAGACGATGATGATCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 1256
 Qy 1048 CTGATGATGAGCGCTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 1107
 Db 1257 CCGCATGATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1316
 Qy 1108 CCGGTGAGTGTGTATACCGGGGGCGAGGGGACGCTGACCTGCTGCTGCTGCTGCTGCT 1167
 Db 1317 CCGCTACAGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1376
 Qy 1168 GACATCCCTGCTGACGAGTGGCTTCTTACCATTTAGCATTAATCTGTGCTGCTGCTG 1227
 Db 1377 GACGTCCAGTGTGACCAAGGCGCCCTGCTCCCATCATATATATCTGTGCTGCTGCTG 1436
 Qy 1228 AATGCTCCCGGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1287
 Db 1437 AACGACCCCTGGAGGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1496
 Qy 1288 GACCGCATGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1347
 Db 1497 GACCGCATGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1556
 Qy 1348 ACCAAAGTGTGACAGCTGAGAGATGCGGGTGTGATGAGACCAAGGGCAAGCCCTCAG 1407
 Db 1557 ACTAAGATGTGAGCTGAGAGATGCGGGTGTGATGAGACCAAGGGCAAGCCCTCAG 1616
 Qy 1408 TATGAGAGGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1464
 Db 1617 TACGAGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1676
 Qy 1465 AAGAGACAGAGCAACTTATCATCATGTCAGAGGAGCACTCAACAAGTCCCTGTGAG 1524
 Db 1677 ATTATGAGCGCTTACCTGTATGCTATGTCTGAGAGAGGTCACAGGATCCCTGTGAG 1736
 Qy 1525 TCTGTGTGATGATCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1584
 Db 1737 TCATGTGAGAGATATAGACTGTGTGGGAGTGTGCTGCTGCTGCTGCTGCTGCTG 1796
 Qy 1585 TGTGTGTGTGTGACACACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1644
 Db 1797 TGTGTGTGTGTGACACACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1856
 Qy 1645 CCGAGTGTGTGCTGAGATGATGAGAGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1704
 Db 1857 AATGATTTGTGCTGAGATGATGAGAGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1916
 Qy 1705 TCGGTCTGTGATGACAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1764
 Db 1917 TCAGTATGAGACAGAGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1976

QY 1765 GCTGCGTCAACTGACCTTTGAGACCTGTGAGAGATGATGGCTGTGCGGCAAT 1824
Db 1977 GCGGGATGCGCCCTGTGCTTTGGGAACTGACAGAGGTGAGGGGAGGTGTGCCGAGC 2036
QY 1825 CAGATCAAGTGTACTTCCCTGTCAGCCAGAGAGGTGCCCCGGATCTACAGAGATGGG 1884
Db 2037 CAGGTATCTGCACTCTCACTGCGGCCAAGAGATG---CCCTGTCTCCCGCTGATCA 2093
QY 1885 GACCAACATGTCGTACAGCTTCACTCAATCAAGAGACGGGACCTGACTCCGACG 1944
Db 2094 GACTGTGTTGGGCTGAGACTACAGCTGAGGTCCAGAGACAGAGAGATTTGTGACG 2153
QY 1945 ACCAGCTTGTCTTCAAAATTCAGAGGTCCACATTCGTGCTGTCTGCTGAGAGT 2004
Db 2154 ACCGAGTTCAAGTTTAAACAATGACAGTCCCAACCACTGTGCTGTGTCAACAGC 2213
QY 2005 CCATACCGCTGCCACCTGTGTAATACCGGCACTGTGCAACCATGACCCCAAGACTG 2064
Db 2214 GCTTTCGCTGCTGCAATGTGTGCAAGTACCGCACTTGTGCACTATGACCCCACTG 2273
QY 2065 TCCCTTCAGAGAGCGCGAGTGAAGCTGCCGAGACTGCCGCCAGCTGTGCGAGTGA 2124
Db 2274 TCCCTTCAGAGAGCGCGAGTGAATATTTCAGAGAGCTGTCCCAAGCTGTGCGCA 2333
QY 2125 AAGATCTGTGCTGCGGTGAGGTGATCAAGCTTACAGCTGAGGCGCAAGAACCTCCC 2184
Db 2334 GAATCTTGAATTCAGTCCGAGGTGAAGCCATACCTTTAAGGCGGAAATCTGCC 2393
QY 2185 CAGCCCAAGTCTGCGGAGCGCTGCTACGAATGATCTCAACATTCAGAGGACGAGCAG 2244
Db 2394 CAGCCCAAGTCTGCGGAGCGCTGCTACGAATGATCTCAACATTCAGAGGACCTCAC 2453
QY 2245 CAGTGTGCGGCTTGTGCTTCAACAGCTCCAGCTGACGTGAGGAGAACCTTTATTCC 2304
Db 2454 CCGGTCCTGCTGTGCTTCAACAGCTCCAGCTGACGTGAGGAGAACCTTTATTCC 2513
QY 2305 TATGAGAGGATGAGATCAACAACCTGCGGCTGAGTTGACAGTCTGTGAGATGGGAC 2364
Db 2514 TATGAGAGGATGAGATCAACAACCTGCGGCTGAGTTGACAGTCTGTGAGATGGGAC 2573
QY 2365 TTCAACATTGACACCCAGCTCAAAATTAAGTTCACTCTCAACAGTGTGAGGCTATG 2424
Db 2574 TTCAACATTGACACCCAGCTCAAAATTAAGTTCACTCTCAACAGTGTGAGGCTATG 2633
QY 2425 GAGAGCTGCGGCTGTGCTTCAAGGCTGACCCAGACTGTCAGATGAGTGTGCGAGG 2484
Db 2634 GAGAGCTGCGGCTGTGCTTCAAGGCTGACCCAGACTGTCAGATGAGTGTGCGAGG 2693
QY 2485 CCAAGGCAAGTGAACCTGTGCGCAGCACTGCTGCCCAAGAGAGCAAGTGTGAGCTG 2544
Db 2694 GAGAGCTGCGGCTGTGCTTCAAGGCTGACCCAGACTGTCAGATGAGTGTGCGAGG 2753
QY 2545 TCTGTGCTGCGGCTGTGCTTCAAGGCTGACCCAGACTGTCAGATGAGTGTGCGAGG 2604
Db 2754 TCTGTGCTGCGGCTGTGCTTCAAGGCTGACCCAGACTGTCAGATGAGTGTGCGAGG 2813
QY 2605 CCGCGGAGAGGAGGAGCAACAGTCACTATCCAGAGGAGAACTGTGAGCTGTGAATTCG 2664
Db 2814 CCGCGGAGAGGAGGAGCAACAGTCACTATCCAGAGGAGAACTGTGAGCTGTGAATTCG 2873
QY 2665 GACATGCGCTCCATGTGCAAGTGTGCTGCGTGAATGAGAGCCCTTTAGTGAATGTTAC 2724
Db 2874 GAGATGCGGCTCCATGTGCAAGTGTGCTGCGTGAATGAGAGCCCTTTAGTGAATGTTAC 2933
QY 2725 ATCCCTGCAAGAGAGTGTGTGAGATGAGGAGGAGGAGCCAGCCAGCAGCATGAGAG 2784
Db 2934 ATCCCTGCAAGAGAGTGTGTGAGATGAGGAGGAGGAGCCAGCCAGCAGCATGAGAG 2993
QY 2785 TTGTGAGATGTGCTGTGCTGTGTGTGCTGAGTGAATTCATGAGCCGCTGTCAAGCTC 2844
Db 2994 CAGATGAGCTGT 3053
QY 2845 TATTAATTCATGACACTGACTCTCTGATCTGAGATCTGAAGCCAGCGGAGGAGCCATGTCCGA 2904

Db 3054 TACACCTTTCGAAACCTTCTGTGTCTGTCACTCAACCAATCCGAGGTCCGAGTCA 3113
QY 2995 GGGAGCCAGAGTACATCAAGGACCAACCTGAATGCGGAGGAGCAAGTGTGTATG 2964
Db 3114 GGCATATGTGTGACATTTACCGGCACTTACCTTGTGGGCTGTGAGACAGCTGTGCTTAC 3173
QY 2965 TTTGAGAGCAGCCCTGTCTTCTTCAAGGAGATCTTCATCTCAATGTGTG---CAAC 3021
Db 3174 CTGGGCAACGAGCTGTGAGTGTCTTACGAGGAGTCAATGATGATGATGATGATGATGATG 3233
QY 3022 ACACATCTTCAGATGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 3081
Db 3234 CCCCATATATCAATGTGCTGTGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3293
QY 3082 ATCCAGAGAGCCTGTCTTCTTCAATGATGAGAGAGCCCAACCTGCGGATTTAGCCA 3141
Db 3294 GTGATGAGCAACCTGTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 3353
QY 3142 GAATGAGCATTTGTGATGAGAAACACCCATGCGCTGTGAGGAGACCCATGAGACTC 3201
Db 3354 GAGTGAAGCATTTGTGATGAGGACACACCCCTGACATCAAGGCTTCAACCTGAGATG 3413
QY 3202 ATACAGAACCCCGATCTCGGCAAGATGAGAGGAGAGGACATCAATATCTGTAG 3261
Db 3414 ATTCAGAGAGCCAGATTCAGATGATGATGATGATGATGATGATGATGATGATGATG 3473
QY 3262 GTTCTGAACGCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3321
Db 3474 GTTCTGAACGCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3533
QY 3322 CAGTGAACCTGACCGAGAGGCGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3381
Db 3534 GCGCTGACATGTGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3593
QY 3382 CTGCTATCTTCAACAGACCACTTCACTATCTTCAACAGGAGGAGGAGGAGGAGG 3441
Db 3594 TTGCTAATTTCAACAGACCACTTCACTATCTTCAACAGGAGGAGGAGGAGGAGG 3653
QY 3442 GGTCCCTCAGAAATCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3501
Db 3654 AGCCCTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3713
QY 3502 CTGATCCCGCTGTGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3561
Db 3714 CTGAGCCCTCTGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3770
QY 3562 AAGCTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3621
Db 3771 ACCCTGT 3830
QY 3622 GGCAGGCAACAAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3681
Db 3831 GGCAGGCAACAAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3890
QY 3682 ATTGCCCCGAGCAGCCGCTGAGCTGTGCGGCAATGTGTGAGATTCGAGTGTGCGG 3741
Db 3891 GTCAATCTCAGACACTTGTGAGCCCTGCAACCACTGTGTGAGCACTGTGCGGAGC 3950
QY 3742 CTCTCATCATTTTCACTGTGTGCGGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAG 3801
Db 3951 CTCTCATCATTTTCACTGTGTGCGGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAG 4010
QY 3802 GACTCAGCTGAAGCGGCTGAGATGAGATGAGCAACTGTGAGTCCCTGTGTGCGGCTG 3861
Db 4011 GACTCAGCTGAAGCGGCTGAGATGAGATGAGCAACTGTGAGTCCCTGTGTGCGGCTG 4070
QY 3862 GAGTGAAGAGAGGCTTTGCGAGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3921
Db 4071 GAGTGAAGAGAGGCTTTGCGAGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4130
QY 3922 GATGAGCGGAGATTCGCTTCTGAGCTATGAACTTACACCATGCGGAGTCTGTCCA 3981

QY 137 AGCCCCGAGGGTTTAAATCACCTGTGTGATGAGAGACAGACATTTACTTGG 196
 Db 683 ATGTGATGAGATTTTCAACATTTTGACTGTACACGAGAAACAGGGGCTGTGTATGTG 742
 QY 197 GGGGCGCATTCGATTTTCAAGGCTCCAGCGACCTGAAAGTCTTGTGTAGACGATGAG 256
 Db 743 GGGGTATCATTCGTGTCTTCAAGTTGACTGCGAACCTCACCAATCCAGTGTGCTCAAGA 802
 QY 257 CAGGGCCGAGCAGAGACACCCCAAGTTTACCCAGCCCGCATGTCTCCAGACTGCAATG 316
 Db 803 CAGGGCCAGAAAGAGACAAAGGCTTGTACCCACCCCTCATTTGTACAGCCCTGCAATG 862
 QY 317 AGCCCCGAGCAGACCAAGATGTGCAAAAGATGCTCTCATAGACTCAAGAGAGACA 376
 Db 863 AAGTGTCTTACCTACCAACATGTCAAACTACTGATCATTTGACTTACTCTGAGAAATC 922
 QY 377 GGGTGAATGCTGTGGAGGCTGTACCAAGGATCTTGCAAGTGTGTGAGGCTGAGAGCC 436
 Db 923 GCGTGTGGCTGTGGAGGCTGTACCAAGGATCTTGCAAGTGTGTGAGGCTGAGAGCC 982
 QY 437 TCTTCAAGCTGGGGAGGCTTATCATAGAGAGAGACATCTGTGAGGTGTCAAGAGA 496
 Db 983 TCTTCAAGCTGGGGAGGCTTATCATAGAGAGAGACATCTGTGAGGTGTCAAGAGA 1042
 QY 497 GCGGCTCAGTCTTGTGAGTATGCTCTCTACAGCAACCTGATGACAAAGCTGTTCATG 556
 Db 1043 CAGGCAACATGTATGT 1102
 QY 557 CCAGGAGATGTGATGGAGAGCCGAGATTTTTCACCAATCTCCAGCCGAGAACTGACCA 616
 Db 1103 GCATGTGTGTGATGGAGAGAGATTAATCTTCTTCTGTGTGTGTGTGTGTGTGTGTGT 1162
 QY 617 AGAATCTGAGAGGCGATGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 676
 Db 1163 GTGACCTGT 1222
 QY 677 TGATTAAGATCCCTTGGACACTTCAACATTCATCCCTGACTTGTATTTACTATGTCT 736
 Db 1223 TCATCAAGATTCCTCTGTGACACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1282
 QY 737 ATGTTTTAGAGTGGCACTTGTGTACTTTTGAACCTTCAACCTGAGATGTGTGTGTGT 796
 Db 1283 ATGTCTTGT 1342
 QY 797 CACCAAGCTCA---CAACCAAGAGAGAGTGTATCATTCAGTGTGTGTGTGTGTGTGT 853
 Db 1343 GCATGGCCATTAATTCAGCTGTGAGACCTTCTTATACCTCAAGATTTGTGTGTGTGTGT 1402
 QY 854 AGGAGGACAGAGCCTTCAACCTGTATGTAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 913
 Db 1403 AGGATGACCCCAAGTTCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1462
 QY 914 TGAATACCGCTGT 973
 Db 1463 TGAATATCGCTTCTGTGAGGACCTTACCTTGTCAAGGAGGAGCTGTGTGTGTGTGTGT 1522
 QY 974 CCGTTGAGT 1033
 Db 1523 CCGTTCAACATGAGAGGAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1582
 QY 1034 GGAATAATGAAATCCCTGT 1093
 Db 1583 AGTACACACACCCCTGT 1642
 QY 1094 ACCGATTAAGAGAGGCTGT 1153
 Db 1643 TGAATATCAAGAGAGGCTGT 1702
 QY 1154 GGGTCAAGGTGAGAGACATCCCTGT 1213
 Db 1703 GGGT 1762
 QY 1214 GTGGCTGT 1273

Db 1763 GGGCTGT 1822
 QY 1274 TCACGAGAGACAGAGACCGCATGACGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1333
 Db 1823 ATACCAAGAGAGAGACCGCATGACGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1882
 QY 1334 TGGCTTGT 1393
 Db 1883 TGGT 1942
 QY 1394 GCAAGCCCTCCAGTATGAGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1450
 Db 1943 ATGT 2002
 QY 1451 ATATGT 2062
 Db 2003 ACATGT 2062
 QY 1511 GAGT 1570
 Db 2063 GGGT 2122
 QY 1571 ACCCCACTGT 1630
 Db 2123 ATCTGT 2182
 QY 1631 GGT 1690
 Db 2183 GGGT 2242
 QY 1691 ATCCCAACATATCTCCGT 1750
 Db 2243 ACCCCACAGATCTGT 2302
 QY 1751 TCCCGAGCTGT 1810
 Db 2303 CTCCCAACCTGT 2362
 QY 1811 TGT 1870
 Db 2363 AGGTATGT 2419
 QY 1871 TACAGAGATGT 1930
 Db 2420 TCCCTGT 2479
 QY 1931 TGAATTCGCGAGCAGAGCTTGT 1990
 Db 2480 AGATCTTGT 2539
 QY 1991 CTTGT 2050
 Db 2540 CTTGT 2599
 QY 2051 ACCCCAGAGCTGT 2110
 Db 2600 ACCCCAGAGCTGT 2659
 QY 2111 TGT 2170
 Db 2660 TGT 2719
 QY 2171 CCAAGAGCTGT 2230
 Db 2720 CCGAAGCTGT 2779
 QY 2231 AGGAG 2290
 Db 2780 AAGGGAG 2839
 QY 2291 ACACTTTATTCATAG 2350

Dp	2840	ACAGCTGTCACCAATATATATGGAAATGGAACATACAGAAACCTTAGCAAGTGGACCTTTCGTATG	2893
Qy	2351	TGTGGAAATGGGCACTTCAACATTTGACAAACCAAGCTCAGATTAAGTTCACTCTTAACAAGT	2410
Dp	2900	TATGGAATATGGAACTTCAATTAATTAACAACCTTCAGAGACCTGAAGATCAATCTTAACAAGT	2959
Qy	2411	GTGAGACCATCGCTAGAGAGCTGCGGAGCTGTGCTCAAGAGCTGACCCAGACTTCCCATGTG	2470
Dp	2960	GTGACAGCCCAACGGGAAGCTGTGGTCTCTGCTCAAGGCTGACCAACAAGTTCAAGTGTG	3019
Qy	2471	GCTGTGTCACAGGCCCCAGGCCCAGTGTACACCTTCCGACGACTTGCCCTTCGCCAGAGACC	2530
Dp	3020	GCTGGTGAAGTGGTAGAGCGCAGATGTACCTCTCCACACAGACTGTGCCCAAGCACTTCTTAGCC	3079
Qy	2531	AGTGGCTGAGACTGTCTGATGACCAAAAGCAAGTGCACAAACCCCGCATCAAGAGATTA	2590
Dp	3080	CCTGGCTTGACTGGTCCAGGCCAACATGTCAAGTGTTCACACCCCAATACAGAGATTT	3139
Qy	2551	TCCCGGTGACAGGCCCCCGGGAAAGGGGACACCAAGTCACTATACGAGGGGAAGAACTGG	2650
Dp	3140	TGAACATATCAGAGACCACTTAAGAGAGGAGACTGTGTACCATCAATGAGCGTGAACCTGG	3199
Qy	2651	GCTGTGAATTTCCGACATATCCGCTCCCATATGCAAGTGTCTGCGGTGAGTGCACCCCTT	2710
Dp	3200	GCTTGAACCTTCGTGAGATGTCTCACCATGTGACAGGTGGCTGAGTGGCCCTGCACACTTA	3259
Qy	2711	TAGTGAATGTTACATCCCTTCGACAAACAGATCGTGTGTGAATATGGGGAGGGCCAAAGCCA	2770
Dp	3260	TCCACGGGGAATACATCAATCGCTAGACAGATCGTGTGAATATGGGCATGCGCTATATAG	3319
Qy	2771	GCCAGCATGACAGGCTTCGTGAGATCTGCGGTGCTGTGTGTCTCGGCTGAAATTACATGAGCC	2830
Dp	3320	GTACACATCTGGGCGCTGTGGCGCTGTGCATTTGGGGAAATGCAAGCAGAGTTCAATGACCA	3379
Qy	2831	GGTCTTCACACTCTATTAATCTTCATGATGACATGACTCTCTCAGATCTGAAAGGCCACGGGG	2890
Dp	3380	AGTCCACACAGATATACTTGTGTGAATCTTCTGTGTGTCACTCAGCCCGATCCGGG	3439
Qy	2891	GAGCCATGTCCGAGAGGAGCCCAAGTGAACCATCACAGGACCAACCTGAATGACCGGAAGCA	2950
Dp	3440	GACCAAGATCAGAGAGTACATGTGTGACCACTACAGGCGCAATTACTTGTGTCTGGAGACA	3499
Qy	2951	ACGTGTGTGTATGTTTGGAAAGAGAGCCGCTGTCTTCCACAGGGATCTCCATCTCAACA	3010
Dp	3500	GTTGTGCAGTATCTGTGGCAATACGACTGTGAATTTATAGGAGGTCTAATGATATGGA	3559
Qy	3011	TTGTCTGCACA - - CCAATCTCCATGTAGAGTCTTAGAGATGAAAGTGTGGTGCAGG	3067
Dp	3560	TTGTATGTGTTCACCCCATCATCATCATGTAGACGTAGACAGTCCCTGTCTCCGTGAGTG	3619
Qy	3068	TGGAAGAGGCGCAAGATCCACACAGAGACCTGTGCTTCAATATGTGGAAACCCCAACATCG	3127
Dp	3620	TGCAAGAGCCCGGGTGTATGACAGTCTGTGCACTTGAATATATAGATATCCACAGGTTCC	3679
Qy	3128	TGCGGATTTGAGCCAGATATGAGCATTTGTCAGTGAAGAACACACCCATGCGGTATGGGGGA	3187
Dp	3680	AACGATTTAGGCAAGATGTAGATATCACTAATGGGGACACACCCCTTAACATCAAGGCT	3739
Qy	3188	CCCAAGCTGAACTCAATACAGAAACCCCGAGATCCGTGGCAACATATGAAGAGGAAGACACA	3247
Dp	3740	TCAATTTGATGTCTTACAGAGCCCAAGGTTCCAGTCAAAATTTATGAGCAAAAGATCTG	3799
Qy	3248	TCAATATCTGTAGAGTTCTGAACGCTACTGAGATGACTGTTCAGAGCGCGGCTTCGCTC	3307
Dp	3800	TCAATGTATGACAGATGTGTAACACACACACCCCTCACTGCTGCACTCTCTGTGACCA	3859
Qy	3308	TGGGTCTGACCAACAGTCAAGCTGACCGAGAGGCCCGAGAGTTTGGCTTCACTCTGG	3367
Dp	3860	GTGACTACCGCCCAAGTCTGACACATGTGGAAGCGGCAGATGAGTTTGGATTTTCTTTTA	3919
Qy	3368	ACAACGTCCAGTCTCTGCTCATCTCTCAACAAAGACCAATTCACTTAATCCCAACCCGG	3427
Dp	3920	ACAATGTTCAATCTTACTCATCTATACACACACCAAGTTCATCTACTACCCCAACCCA	3979

QY	3428	TGTTTAGGCGCTTGGTTCCTCGAAGATCCTGAGAGCTCAAGCGCTGGACGCCCATCATCC	3487
Db	3880	CGTTTGAACGTCTAGCGCCCACTGGAAATCTTGGATCAAGAGCCAGGCTCACCCATCATCC	4039
QY	3488	TAAAGGCGAAGAACCTTATCCCGCTGTGGCTGGGGCAACGTGAAGCTGAACCTACATG	3547
Db	4040	TGAAGGCGAAAAATCTGTCTCTCTCTCTGGAGG---GGCCAAATCAACTACAG	4096
QY	3548	TGCTGTGTGGGAGAAACCGCTGACCGGTGACCGGTCAATGTCACATGCTCTGGAGT	3607
Db	4097	TAAATGATGGAGAACCTTGTACAGTCACTGTGTGTGAGACACACCTGCTTTGTAAAC	4156
QY	3608	CCCCCAACCTTATGCGAGGCACAAAGTGAATGGCCCGCTGTCCGTGGAGATGACTCCC	3667
Db	4157	CTCCCAACCTCACAGGGCAGACACAAAGTCAATGGTTCAAGTGGGGGATGGATGTTCCAC	4216
QY	3668	CGGGAGTGGGTATGATTTGCCCGGAGACCGCGCTACGCTGCCCGCATGTCAGCACTCG	3727
Db	4217	CTGGCTCCGTGAGCGTATCTCCGACAGCCTGTGTGAACCTGTCCAGCATATCAGCATCG	4276
QY	3728	CAGTGGCTGGCGGCTCTCATATTTTCATGTGGCCGTCTCATTTGCTTATTAACGCA	3787
Db	4277	CAGCTGGTGGAGCGCTCTTTCATCTGTCATCATTTGTCATCATGCTTACAAAGGCA	4386
QY	3788	AGTCCCGGAAAGTGAACCTCACGCTGAAGGGGCTGCGAGATGACGATGAGACAACTGGAGT	3847
Db	4337	AGTTAAGGAATATGACTTCACTCTCAAGCGCTCCAAATGCAATGGAACACTGGAGT	4396
QY	3848	CCCGTGTGGCCCTGAGATGCAAGGAAGCCTTTGCGAGCTGACAGCGAOCATCCATAGC	3907
Db	4397	CCAGGGTGGCACTGAGTGCAGAGGAACCTTTGGGAGCTTCAGACAGACATCATAGC	4456
QY	3908	TGACCAATGACCTGGATGGAGCGGGATTCGGTTCCTGGAATATAGAATTACACATAGC	3957
Db	4457	TACCAATGACTTGGATCGATCAGAAATCCCTTACCTGGACTACCGTACCTATGCTCAGA	4516
QY	3968	GAGTGTCTGTTCCAGGAATTGAAGAACCAACCTGTCTCCGGAGCCTTGAAGTCCGGCT	4027
Db	4517	GAGTCTGTTCCAGGATTGAGAGCAACCTGTCTGCGGGAACTGGAGGTACAGGGAA	4576
QY	4028	ACCGGAGAGGCGGTGGAGAAAGGCGCTGAAGCTCTCGCCAGCTATCAACAACAAG	4087
Db	4577	ATGACACAGCACAGTGGAGAAAGCCTGAATCTTTCGCCAGCTTATCAACAACAAG	4636
QY	4088	TGTTCTGTCTCTCTCATCCGACGCTTGAATCCAGCGTAACTTCCATGCGCAAC	4147
Db	4637	TGTTCTGTCTAACCCTTATCCGTCACTGGAATCAAGCGAGCGCTTCCATGCGAAGC	4696
QY	4148	GTGGCAACGTGGGCTCACTATCACTGACCGGTGCGAGAGCAACTGGAAGAACCCACTG	4207
Db	4697	GTGGAAAGTGGCTCTCTCATCTGACAGGCTTCAAGGTGGCTGAATATGCCACTG	4756
QY	4208	ATGTGTGAAGACAGCTGCTGGCGCACTCATTTGAAGAACTGGAAGACAAACAAC	4267
Db	4757	ATGTCTTCAGACAGCTGCTCTCTGACCTCAATTTGAAGAACTGGAAGAAACAACAAC	4816
QY	4268	CTAAGCTGCTGCTGAGGAGCATGAGTCAGGTGCTGAAGAGATCTGACCAATTTGTTTA	4327
Db	4817	CCAGCTGCTCTCTCGCAGACATGATGTGTGGCCGAGAAAGATCTGACTAATCTGTTTG	4876
QY	4328	CTTTCCTCTCTCAAGTTCCTCAAGAGATGTGTGCGGAGCCCTCTTCTCCTGTTC	4387
Db	4877	CTTTCCTCTCAACAAGTTCCTGAAGAGATGTGTGCGGAAACAACCTTCAATGTATACT	4936
QY	4388	GTGCATTAAGACACAGATGAGAGAGGCCCATTTGACGCATACAGGCGACAGCCGCT	4447
Db	4937	GTGCATTAAGACACAGATGAGAGAGGCCCATTTGACGCATTAACGTGTGATCTGTA	4996
QY	4448	ACTCCTTAGCGAGAGCAAGCTCATTCGCCACAGATTAAGTACTACAAAACCTGTCTCGA	4507
Db	4997	ACTCCTTAGAGAGACAAAGCTCATTCGGCAGCAGATGAGATTAAGACTGTGATCTGTA	5056

4508 GCTGTGTCAGCCAGACAAATGCAACAGCCCGAGTCCAGTAAAGATCCTCAACTGTG 4567
 5057 ACTGTGTCAACCCCTGACAAATAGAACAGCCAGAGATCCAGTGAAGTACTAAACTGTG 5116
 4568 ACACCAATCACTCAGGTCAAGAGAAATTTCTGATGTCACATCTTCAAGATGTGCTGCT 4627
 5117 ACACCAATCACTCAGGTCAAGAGAAATTTCTGATGTCACATCTTCAAGATGTGCTGCT 5176
 4628 CCCACCGGCCCCAAGCTGACAGATGATGATGTCGAGTGGCGCAAGAAAGTGGGCAAGA 4687
 5177 CCCAGCGGCCAAGGCTGTGACATGATGATGTCGAGTGGCGCAAGAAAGTGGGCAAGA 5236
 4688 TGATCTTTCAGAGATGAAGACATCAACCAAGATTGAAATGATTGGAGCGACTGAACA 4747
 5237 TGTGTGTCAGAGACGAAGACATCAACCAAGATTGAAATGATTGGAGCGACTGAACA 5296
 4748 CACTGCCCCCACTACAGGTGACAGATGATGATGTCGAGTGGCGCAAGAAAGTGGGCAAGA 4807
 5297 CACTGATGATTAACAGGTGACAGATGATGATGTCGAGTGGCGCAAGAAAGTGGGCAAGA 5356
 4808 CAGCCTATTAACGACAGTGAACAACTCAACCGTCTCCAGAGCTCAGCAAGTAAATATGAAA 4867
 5357 CCTCCTTAACATCCCTGAGCTGTCAGCAATCTCTCGGACATCCATTAAGCAGATATGACT 5416
 4868 ACATGATCCGCTAACAAGGCGAGCCCGCAGCCTCCGCTCAGGACACTATGATCACTC 4927
 5417 CTTCCTTCAAGTACACAGGCGAGCCCGCAGCCTCCGCTCCGCTCCGCTCCATATCAACC 5476
 4928 CTGACCTGAGAGTGAAGTGAAGTGTGACCTAGTGAAGAAACCAAGACAGGAGACC 4987
 5477 CAGACTTGAGAGGCGGTGTCAAGTTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 5536
 4988 AGAAGAGAGGCGGAGCCGCGGAGCAAGATGCTGTGAATCTTACCTGACCCGACTCTCG 5047
 5537 AGAAGAGAGGCGGAGCCGCGGAGCAAGATGCTGTGAATCTTACCTGACCCGACTCTCG 5596
 5048 CCATTAAGGAGGAGCTGAGAGATTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 5107
 5597 CCATTAAGGAGGAGCTGAGAGATTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 5556
 5108 CACACCGTGGCTCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 5167
 5557 TGACACCGGAGTGTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 5716
 5168 AGGCTGATTAACAGGATTCATGATGATGATGATGATGATGATGATGATGATGATGATG 5227
 5717 AGGAGAGACAGACAGATTCATGATGATGATGATGATGATGATGATGATGATGATGATG 5776
 5228 TGCCCTGAGGTTTGGGTCAACATGATGATGATGATGATGATGATGATGATGATGATG 5287
 5777 TTCCACTTCGTTTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 5836
 5288 AGAAGAGATCAAGAGGCTGCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 5347
 5837 AGGAGAGATCAAGAGGCTGCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 5896
 5348 CCAGGTCAAGACAGGCTGAGGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 5407
 5897 CCAATGAGAGACAGGCTGAGGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 5956
 5408 ACATGCCCCAGTCAAGATGAGTGTGAGGATTTACTGACATGAGGAGATGAGTGTGAG 5467
 5957 ATATGCCCCAGTCAAGATGAGTGTGAGGATTTACTGACATGAGGAGATGAGTGTGAG 6016
 5468 CCATCAGGACCAAGACATGAGGATGAGTGTGAGGATGAGTGTGAGGATGAGTGTGAGG 5527
 6017 CCATTAAGACCAAGATGAGTGTGAGGATGAGTGTGAGGATGAGTGTGAGGATGAGTGTG 6076
 5528 AGTTCAACACCATGAGTGTGAGTGTGAGGATGAGTGTGAGGATGAGTGTGAGGATGAG 5587
 6077 AGTTCAATATGAGTGTGAGGATGAGTGTGAGGATGAGTGTGAGGATGAGTGTGAGG 6136
 5588 AGATCTTGAAGCTGTGAGACAGATGAGTGTGAGGATGAGTGTGAGGATGAGTGTGAGG 5647

Db 6137 ACCTATGAGGAGCACTAGAGCAGATGAACAGCCCGCAGCAAGCACTGAGCTCAAGG 6196
 QY 5648 TAGAACCACTCTAACCTCCTAGACTTGAACAGTGAATGA 5691
 Db 6197 TGGACATCTCATCAACCCCATGTCATGAGAGCTGAAGGAA 6240
 RESULT 7
 LOCUS D86949 6730 bp mRNA linear ROD 31-MAR-2000
 DEFINITION Mus musculus mRNA for plexin 2, complete cds.
 ACCESSION D86949
 VERSION D86949.1 GI:1655431
 KEYWORDS plexin 2.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (sites)
 Kameyama, T., Murakami, Y., Suto, F., Kawakami, A., Takagi, S.,
 Hirata, T. and Fujisawa, H.
 Identification of plexin family molecules in mice
 J. Cell Biol. 148 (6), 1283-1293 (2000)
 MEDLINE 96400270
 PUBMED 8806646
 2 (sites)
 Shimizu, M., Murakami, Y., Suto, F. and Fujisawa, H.
 Determination of cell adhesion sites of neuropilin-1
 J. Cell Biol. 148 (6), 1283-1293 (2000)
 MEDLINE 20191894
 PUBMED 10725340
 3 (bases 1 to 6730)
 Fujisawa, H.
 Direct Submission
 Submitted (07-AUG-1996) Hajime Fujisawa, Nagoya University Graduate
 School of Science, Division of Biological Science, Furo-cho,
 Chikusa-ku, Nagoya, Aichi 464-01, Japan
 (E-mail: fujisawa@bio.nagoya-u.ac.jp, Tel:052-789-2978,
 Fax:052-789-2979)
 FEATURES
 source
 CDS
 1..6730
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /tissue_type="brain"
 /dev_stage="E17.5"
 580..6234
 /codon_start=1
 /product="Plexin 2"
 /protein_id="BA13189.1"
 /db_xref="GI:1655432"
 /translation="MADNLISVLLISVAMLLIARGTGMPQSTFSEHSDMTFENHLL
 VHRGATVYGAIRNRYVLTGNTLTQVAKHTGPEHNDACYPILVPCSSVLTLLTN
 VNKLLIDYSEKRLACSLKYGVCKLRDLDFILVPSKHKHYLSVAKTGMV
 YIVREGEDEGLFETGAVDGKDYFPLISSKRLRPDESSAMLEHLSDDVSILKI
 PSDTLAVSHSDIPYIVGFAGGVYFVLPPEPDLQAINASGLFETSLIVRCKD
 PKFHSYVSLPFGCTRAVEYRLQAALAPAGBALQAFNISDSDVLFALFSKGGK
 OYHEPDDSLCAPIRALINLOIKERLQSCYHGSNLEIMNLGIDVCTRAPYVID
 NPGCILDINPLGSGTPVEGLTLTTSRRLSVASVYVNGSVFVGTGSKLKKIRA
 DEPRHGYQYEMVSYFKDGSPIILDMAESIINOLYLWMSRQUTRVPESGEYTTG
 ECLSDGPHGCMCAHNKCSRDKQKRMENRRAISQCSLEFHNHSISYSDHSR
 LLTLVNDAPNLSEIACAFNLTVEBQVGSQVLCISPEKADVPVLPDQDMGLE
 LQLRKETGKI FVSTEFKYNCAHQCLSCVNSAFRCHWKYRNLCHDFTCSFQ
 GRINVEDCPDLPVTEILLIPVEKQPLITLARNLPQPSQGRYEVLSIQGAVRV
 PALNRSSVOCNSYQYDMDISLNAVDAVWNGNFIIDNPDLVHLYKCAQR
 ESKGLKADHKFECGSGRGLCHGCTSPMDMSHNVCNIPDITELLTV
 TSPGRIRVITVHGNLTIDRSKLAHNVAGVCTPIRGYITIAEOIVCMGAVIG
 GSGPYRCLCEKCEKRPMTKSHCOYTPNPSVLSIPRGRSGGTWVTLTGAVAG
 SSVAVYLGNTQGERFGKSMETVCSPPSSGLPVPVSVYDPRADSSIQFEYID
 PRVQRIPEWMSITSGHTPLITGFLNDVIGPVRVFNKGSVAVCTVATTTCL
 APSLSDYRPGIDTVRDEFGFLNNVQSLIYNDFIYIPNPTFELSLPTGLDQ
 KFGSPILLKKNLCPASGAGAKLVYTWIGETPCTVVSETOLLCEPNLITGOKWV

HYGMVPSPPSVISDSLLTLPALISIAAGSLLIIVIIIVLAIYKRENDLTIK
RLQMDNLESFVLECKEAPAFLEQTDINELISDLDRGIDYLYRYAMVLPGE
LHPVLELEVQNGQOHEKAKLFLPILFRTLELQSRSMKDRNVA
LIMTGLQRLLEVATDVLLKOLISDLIDKMLENNKPKLRLRTEVAEMLMNPAFL
HFLKEAGEPLFMYLCAIKOOMEKPIIDATGEARYSLSDKLIROQIEKTLINC
KVPDMENBEIPVKLINDTLIOVEKILIDVYKNRYPVSPRAVIDLEMRGRIR
VYLQEDITTKRESDEPKRLMTHYOVSDRSVALYPCOTSXNIPASASISRTISR
YDSFRYTESPDLRSRVPMITPDLESGRVHLVKNHDHEDQKDRGSRMVEIYL
TRLATKGLQKFDVDFEPLTFSTVHRGSALPLAKYDFDLSDADHSHIDTVH
TWKSNCLPLRFWNVYIKNOQVFDIHKISIDACLSVAQFMSQSTSEHRLKXDP
SNKLIYADIPSKWVERVYADIKLPALISDQMNAYLAQOSRHHATEFMMALNE
IYSVYSKXSEELIHALDEQEARQRLAYKVEHLINMISIES"

ORIGIN

Query Match 48.8%; Score 2778.4; DB 10; Length 6730;
Best Local Similarity 69.3%; Pred. No. 0;
Matches 3858; Conservative 0; Mismatches 1691; Indels 15; Gaps 5;

QY 137 AGCCCGCAGGAGTTTCATCACTGTGTGATGAGAGACAGACATTACTTG 196
DB 683 ATGTGACTGGAATTCACCACTTGTACTGACCGAAGACAGGGCTGTATG 742
QY 197 GGGCCGTCAATGGATTACAGCTCTCCAGCACTGAGTCTTGTGACGATGA 256
DB 743 GGGCTATCAATGATGTCTCAAGTTGACTGGCACTCCATCCAGGTGGCTCA 802
QY 257 CAGGGCCGAGAGAGCAACCCCAAGTGTACCAACCCCGCATCGTCCAGACTG 316
DB 803 CAGGGCCAGAGAGAGCAAGGCTTGTACCCACCTCTATTGTACAGCCCTG 862
QY 317 AGCCCTGACCAACCAACATGTACAGAGATGCTCTCATAGACTACAGAGAA 376
DB 863 AAGTGTACACTCAACCAACATGTACAGAACTGTATCATTACTGTGAATC 922
QY 377 GGTGATTCCTCTGGAGAGCTGTACAGAGCATGTGAGTGGCTGGAGAGCC 436
DB 923 GCTGTGCTGGAGAGAGCTGTACAGAGGTTGTGCAAGTCTCTGCACTAGT 982
QY 437 TCTTCAAGCTGGGGAGCTTATCATAGAGAGAGCATATGTGTCAAGTGTCA 496
DB 983 TCTTCACTGTGGAGAGCTTCCACAGAGAGACACTACTGTCTCAAGTGTCA 1042
QY 497 GGGGCTCATCTTGGAGATGTGTCTCTACAGCACTGTATGACAGGCTTCAT 556
DB 1043 CAGGCAACATGTATGTGTGTGTGCTGTAGGGGAGAAATGACAGCTTTTAT 1102
QY 557 CCAAGGAGTGTAGAGAGCCGAGATTTTCCACCATCTCCAGCCGAGAACT 616
DB 1103 GCACTGTGTGATGGAGAGCAAGATTACTTCCCTACTCTGTCCAGCCG 1162
QY 617 AGAAGCTGTAGAGAGAGATGTGTGGCGTAGCTCTTCATGATGATTCGT 676
DB 1163 GTAGCCCTGAGTTCAGCAATGTGTGACTATGAGCTCCAGATGATTTTCT 1222
QY 677 TGATTAAGATCCCTTGGACACCTTCACCATCACTCCCTGACTTGTATCT 736
DB 1223 TATTAAGATCCCTTGGACACCTTCAGCTGTCTCTCATGAGAAATGTG 1282
QY 737 ATGCTTTAGAGAGTGTGCACTTTTGTACCTTCCAACTGTAGATGTGTCT 796
DB 1283 ATGCTTTAGAGAGTGTGCACTTTTGTACCTTCCAACTGTAGATGTGTCT 1342
QY 797 CACCAAGGCTCA---CCACCAAGAGAGAGTGTATCATCAAGCTGTAGAG 853
DB 1343 GCAATGGCATCAATTAAGCTGAGAGACCTTCTATACCTCAAGAAATGTG 1402
QY 854 AGAAGAGACAGGCTTCAATCTCTATGTAGAGTGTGCTGTGAGCGAGTGG 913
DB 1403 AGATAGACCCCAAGTTCACTCTATGTGTCTCTGCTTTTGTGCTCACA 1462
QY 914 TGAGATTCAGGCTGTGAGAGGCTCACTGTCCAAAGGAGGAGCGGTGTGG 973
DB 1463 TGAATATCGCTTCTGAGGAGCTTACTTGTCAAGAGCGAGAGAGCTTA 1522

QY 974 CCTTGAAGTCATCCAGATGATGACCTGTCTTACCGGTCTTTCTCCAGAGCC 1033
DB 1523 CCTTCAACATCAACAGAGAGAGATGCTCTGTGTGCAATCTTTTCCAGAG 1582
QY 1034 GAAATATGAAATCCCTGATGAGTGGGCTGTGCACTTCTCATCTTGAAGCA 1093
DB 1583 AGTACACCAACCCCTGATGATCTCGCTCTGTGTCTTCCCATCCGAGCAT 1642
QY 1094 ACCGATTAAGAGCGCTGCAATCTTGTACCGGGGCGAGAGGCACTGTG 1153
DB 1643 TGCATATCAAGAGCGGTGCAATCTGTGTACCAAGAGAGGCACTTGA 1702
QY 1154 GGCCTAAGTGAAGAGCATCCCTGAGAGAGTGCCTCTTAACTTGAAGATCT 1213
DB 1703 GGTCTGTGGAGAGAGATGTGCAATGACCAAGCGCTGTCTCCAAATGAT 1762
QY 1214 GTGCTGTGACATGAATGCTCCCTGTGAGAGTGTCCGACATGTGTGAT 1273
DB 1763 GCGGCTGTGACATGAACAGCTCTGTGAGGCTCACTCTGTGAGAGAG 1822
QY 1274 TCAAGAGAGCAGAGACCGCATGACGTGTGTATGCTATATGCTCAAG 1333
DB 1823 ATACACACAGAGAGAGCGCTTACCTGTGTGCTCTATGTATGATGAT 1882
QY 1334 TGGCTTTGTGGCACCAAAAGTGGCAAGCTGAAGAGATCCGGTGTATG 1393
DB 1883 TGGTTTGTGGGAGATGAAGTGGCAAGCTGAAGAGATCCAGCTGATGT 1942
QY 1394 GCAAGCGCTCCAGTATGAGAGCGTGAAGTGTGT---GGACCCGCGCC 1450
DB 1943 ATGTTGGGTGCTGATGAGATGTCTGTGTCAAGATGAGAGAGCCAT 2002
QY 1451 ATATGCTTCTCCAGAGACAGAGCACTTACATCATGTGAGAGAGCTCA 1510
DB 2003 ACATGGCTTCTCATCACTACATCACTATATGATGTGTGAGAGAGAG 2062
QY 1511 GAGTCTGTGAGTCTGTGTGTATGATGATGAGAGCTGGGAGAGTGTG 1570
DB 2063 GGTGCTGTGTGATCATGTGAACAGTATCAACTGTGAGAGTGTCTA 2122
QY 1571 ACCCCCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1630
DB 2123 ATCTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2182
QY 1631 GGTCAAGAGAGCCCGAGGTTTGTGTGTGTGTGTGTGTGTGTGTGTGT 1690
DB 2183 GGGCTGTGAGAGCAATGATGTGTGTGTGTGTGTGTGTGTGTGTGT 2242
QY 1691 ATCCCAATATCTCCGTCTCTAGTACAGAGTGTGTGTGTGTGTGTGTGT 1750
DB 2243 ACCCAACAGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2302
QY 1751 TCCCGAGCTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1810
DB 2303 CTCCCACTCTGTGAAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2362
QY 1811 TGT 1870
DB 2363 AGTATGT 2419
QY 1871 TCACAGAGATGAGAGACCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1930
DB 2420 TCCCTGT 2479
QY 1931 TGACCTTGT 1990
DB 2480 AGATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2539
QY 1991 CTTGCTGT 2050
DB 2540 CTTGT 2599

QY	2051	ACCCCAAGACCTGTCTCCCTCCAGAGAAAGGCCAGTGAAGCTCCCGAGACCTAGCCCCAGC	2115
Db	2600	ACCCTCACTACTGTTCCTTCTCCAGAGAGGACAGATCAATGTTTCAGAGGACTGTCTCCAGC	2659
QY	2111	TGCTGCAAGTGGACAAGATCCTGTGTGCCCTGAGAGTGATCAAGCTTATCAAGCTGAAGG	2170
Db	2660	TCTGTCCACAGAGAGAGATTTCTGATTCCTCAGTTGGGGAAAGTAAAAACAATCCCTTAAAGG	2719
QY	2171	CCAAAGAAGCTCCCAAGGCCCAAGTCTGGGAGCGTGGCTAGAGATGCACTCTCAACATTC	2230
Db	2720	CCCCAAGACTGCCCCCAGCCCAAGTCTGGCCACAGCAGAGCTACAGAGTGTGTGCTCAGCAATTC	2779
QY	2231	AGGGCAGCGAGCAGCAGAGTGCCTGCCCTGACCTTCAACAGCTCCAGGCTACAGTCCGAGA	2290
Db	2780	AAAGGGGCTGTCCACCGGGTCTCCCTGCCCTGCGGTTTCAACAGTTCCAGTGTGAGTGGCCAAA	2839
QY	2291	ACACCTCTTATTCCTATGAGAGGATGAGAGATCAACAACTTCGCTCGGTGAGTTCACATCG	2350
Db	2840	ACACCTCGTACCAAGTATGATGGATGTGACATCAACACAACTAGCAGTGAATTTCCTGTAG	2899
QY	2351	TGTGAATATGGGCACTTCAACATTTACAAACCAGCTCAGATTAAGTTACCTCTACAAAT	2410
Db	2900	TATGATATGGCACTTCAATTTATGCAACCTTCAGAGACCTGAAGTACATCTTACAAAT	2959
QY	2411	GTGAGCCCAATGCTGAGAGCTGCGGGCTGTGCTCAAGGCTGACCCAGACTTGCATGTG	2470
Db	2960	GTGACGCCCAAGCGGGAAAGCTGTGGTCTCTGCTCAAGGCTGACCAACAGTTCCAGTGTG	3019
QY	2471	GCTGTGTGCAAGGGCCCAAGCCAGTGCACCCGTGGCGCAGACATGCGCTGCCAGAGAGCC	2530
Db	3020	GCTGTGTGAGTGTGTGAGCGCAGATGTACCTTCAACACAGCACTTGCCCAAGCATTTCTTAGCC	3079
QY	2531	AGTGGCTGAGAGCTGTCTGTGTCCTCAAAAGCAAGTGCACAAACCCCGCATCAAGAGATAA	2590
Db	3080	CCTGGCTTGAATGTGTCCAGCAACAATGTCAAGTGTCCAAACCCCAATACAGAGATTT	3139
QY	2591	TCCCGGTGACAGGCCCCCGGGAAAGGGGACACAAAGTCACTATCCAGAGGGAGAACTGG	2650
Db	3140	TGACAGATATCAGGACCACTGAAGAGGGAGATCTGTGACCATCAATGGCTGAAACCTGG	3199
QY	2651	GCCGTGAATTTCCGGAACATGCGCTCCCATCTCAAGGTTGCTGGCGTGGAGTGCAGCCCTT	2710
Db	3200	GCTTGGACTTCTGTGAAGTTGTCTCAACATGTGAGAGTGGCTGGAGTGGCCCTGCAGACCTA	3259
QY	2711	TAGTGATGTTAATCATCCCTCGAABAACAGATCGTGTGTGAGATGGGGGAGGCCAAGCCCA	2770
Db	3260	TCCCAAGGGAAATACATCATCGCTGAGCAGATCGTGTGAGATGGGCGCATCCGTTATAG	3319
QY	2771	GCCAGCATGCAAGCGTGGTGAGATCTGCGGTGGCGTGTGTCGGCTGCAATCAATAGGCC	2830
Db	3320	GTACCACTCTGGGCGCTTGGCGCTGTGCATTTGGGAAATGCAGGCGAGATTCATGACCA	3379
QY	2831	GATCCTCAAGACTCTATTAATCTTCATGACATCGACTCTCTCAAGATCTGAAGCCACCGGG	2890
Db	3380	AGTCCCAACAGAGATATCTTTTGTGATTCCTGTGTGCTGTCACTAGCCCAATCCGGG	3439
QY	2891	GGCCCATGTCCGGAGGGACCCAAAGTGAACATCAACAGGACCAACTGAAATCCCGGAGCA	2950
Db	3440	GACCAGATCCAGAGAGTACATGTGACCAATCAACAGGCAATTAATCTTGCTGTGGAGCA	3499
QY	2951	ACGTGTGTGATGTTTGTGAAAGCAGCCCTGTCTTTCACAGGCGAATCTCATCTTACA	3010
Db	3500	GTTGTGGCAGTTCACCTGGGCAATCAGACTGTGAATTTATGGGAGGTCAATGATAGGA	3559
QY	3011	TTGTCTGCAACA--CCAATCTCAATGAGAGTGTGAGATGAAGGTGTGCGTCAAG	3067
Db	3560	TTGATATGTTTCAACCCCAATCATCAATGACCTAGACCAAGTCCGTGTCTCGTGAATG	3619
QY	3068	TGACAGAGGCGCAAGTCCACAGAGCACTGTGCTTTCAGTATGTGAAACCCCAACATCG	3127
Db	3620	TGACAGAGCCCGGGTGAATACAGCTGTGCAGTTGAGATTAATAGATACCCACGCGTCC	3679
QY	3128	TGCGGATTTAGCCAGATGAGACATTTGATGTGAAACAACCCATCGCGGTATAGGGGGA	3187

Db	3680	AACGATATTGAGCCAGAGTGAGATCACTAGTGGGCACAACCCCTTAACCATCAAGGCT	3739
OY	3188	CCCACTTGAGACTTATACAGAACCCCAAGATCCGTGCCAAGATAGAGGGAAAGAGACA	3244
Db	3740	TCAACTTGAGATGATATCAGAGAGCCAGGGCTCCAGTCAAAATTATAGCAAAAGATCTG	3799
OY	3248	TCATATCTGTGAGAGGTTCTGAAACGCTACTGAGATGACTGTGCAGAGCCGCCCTCGCTC	3307
Db	3800	TCAATGTATGACAGCTGTGTAAACAAACACACCCCTCACTGTCTGGACACCTCTGTACCA	3855
OY	3308	TGGGTCTCTGACCAACAAGTCAAGACTGAACCGAGAGCCCGAGAGATTGGCTTATCCTGG	3367
Db	3860	GTGACTACCGCCAGGTCTGGACACTGTGTGAAACGGCCAGATGAATTTGGATTTCCTTTA	3919
OY	3368	ACAACGTCCAGTCCCTGCTCATCTCTCAACAAGACCAATTCACCTATCATCCACCCGG	3422
Db	3920	ACAAATGTTCAATCTTATCTATCTATACACACACCAAGTTATCTATCTACCCACCCA	3979
OY	3428	TGTTTGAAGCCCTTGTGTCCCTCAGAGAAATCCTGAGACTCAAGCCTTGGACCGCCATATCC	3487
Db	3980	CGTTTGAAGTCTCAGCCCCCACTGGAAATCTTGATTCAGAAAGCCAGGCTCAACCCATATCC	4039
OY	3488	TAAAGGCGAAGAACCTGATCCCGCTGTGGTGGGGGCAAGTGAAGCTGAACATCAACTG	3547
Db	4040	TGAAGGCGAATAAATCTGTCTCTCTCTCTCTTGGAG--GGCCAAATCAATPACAG	4098
OY	3548	TGCTGTGTGGGAGAACCCGTGCACCCGTGACCCGTGCAGATGTCCAGTCTCTCGAGT	3607
Db	4097	TAAATGATTGAGAGAACCTTGTATCAGTCACTGTGTCTGAGACACAGTGTGTTGTGAC	4156
OY	3608	CCCCCAACTCATCGGCAAGGCAAAAGTATGCCCCGTGTGCTGGTGGCATGGAATCTCC	3667
Db	4157	CTCCCAACTCACAGGGGACGACAAAGTCAATGTTCACTGTGGGGGAGTGTGTTCTAC	4216
OY	3668	CGGGGATGAGTGTACATTGCCCCCGGACAGCCCGCTCAGGCTGCCCGCATCTGCAGATCG	3722
Db	4217	CTGGCTCCGTAGAGCTATCTCCACAGCCCTTGTGACCTCTGCAGCCATCATCAGATCG	4276
OY	3728	CAGTGGCTGCGGCGCTCTCTCATCATTTTCATGTGGCCGTGCTCATTCCTATTAACGCA	3787
Db	4277	CAGCTGGAGAGGCTCTCTCTTATCATCTGCATCATTTGTCTCTCATCGCTTAACGCCCA	4338
OY	3788	AGTCCCGGAAAGTGAACCTCAAGCTGAAGGGGCTGCAGATGACATGAGCAACCTGGAGT	3844
Db	4337	AGCTTAGGAAATGACCTCACTCAAGCGGCTTCCAAATCCAAATGAGCAACTCTGAGT	4396
OY	3848	CCCGTGTGGCCCTTGAGTGCAGAGAAAGCTTTGCCGAGCTGCAGACGACATCCATGAGC	3907
Db	4397	CGAAGGTGCACTGAGAGTGCAGAGAAAGCTTTTGGGAGCTTCAGACAGATCATGATGAGC	4456
OY	3908	TGACCACTGACCTGATGAGACCCGGATTCGGTTCCTGGATATATAGACTTAACCCATGC	3966
Db	4457	TAAACAGTGAATTGATGATCAGAAATCCCTTACTGTGACTTACGTAACCTATGTGCATGA	4511
OY	3968	GAGTGTGCTTCCAGAAATGAAACCAACCTGTACTCTCCGAGACCTTGAGTCCCGAGCT	4027
Db	4517	GAGTCTGTCTCCAGGATTTGAGGACACACCTGTCTTGGGGGAAGTGGAGTGAACGGAA	4576
OY	4028	ACCAGCAAGAGCGTGTGAGAAAGGCTGAAGCTCTTTCGCCAGCTCATCAACAACAGG	4087
Db	4577	ATGACAGAGCAACAGTGAAGAAACCCCTGAAACCTTTCGCCAGCTTATCAACAACAAG	4636
OY	4088	TGTTCCGCTGCTCTTATATCCGCAACGCTGAGTCCCAAGGCTTCCATGCGCGAC	4144
Db	4637	TGTTCTTGTCTAATCTTCAATCCGTACACCTGGAACATPACAGCGAGCTTCTCCATGCGAAGC	4696
OY	4148	GTGCAACGTGGGCTCACTCATATGACCGTGTGCAAGCAAGCTGAGTACGCTCATG	4207
Db	4697	GTGGGAACGTGGCTCTCTCTCATATGACAGGCTTTCAGGGTTCGCTTAATAATGTCACTG	4756
OY	4208	ATGTGCTGAACACACTGTGGCCGACCTCATTTGACAAAGACTGTGAGCAAGAACCAAC	4266

Db	4757	ATGTCCTCAACGACGTGCTCTCTGACCTCTCATTTGACAGAAACCTGGAGAAACAAGAACCAACC	4818
QY	4268	CTAAGCTGCTGCTCAGAGGAGCATGATCAGTGGGCTGGAGAAATGCTGACCAATTGGTTA	4327
Db	4817	CCAAAGCTGCTTCTCCGACGAGACTGATGTTGTGGCCGAGAAATGCTGACTAAGCTGTTTG	4876
QY	4328	CTTTCCTCTCTACAAAGTTCTCTCAAGAGATGTGCTGGGAGCCCTCTCTCCCTGTTCT	4387
Db	4877	CTTTTCTTCTACACAAAGTTCCTGAAAGAGTGTGGGGGAAACCACTCTTCAAGCTATACT	4936
QY	4388	GTGCATCAAGCAGAGATGAGAAAGGGCCCATTTGACGCATACAGGGGAGAGGCCGCT	4447
Db	4937	GTGCATCAAGCAGAGATGAGAAAAAGCCCCCATTTGACGCTATTACTGTGTAGGCCGAT	4996
QY	4448	ACTCCTTGAGCGAGGACAAAGCTCATCCGACAGAGATTTGACTACAAAACCTGTGCTGA	4507
Db	4997	ACTCCTTGAGTGAAGACAAAGCTCATCCGGACAGAGATGAAATTAAGACTCTGTACTGA	5056
QY	4508	GCTGCTGAGCCCGACAAATGCCAACACCCCGAGGTCCGAGTAAAGATCCTCAACTGTG	4567
Db	5057	ACTGTGTCAACCTTGACAAATGAGAACGCCAGAGATCCGATGAAGTACTTAACGTGTG	5116
QY	4568	ACACCATCACTCAAGTCAAGAGAAAGATTTGATGATCCATCTTCAAGATGTGCTTGTCT	4627
Db	5117	ACACCATCACTCAAGTCAAGAGAAAGATCCTCGATGCCGTATTAAAGATGTCCCTACT	5176
QY	4628	CCCAACGGGCCAAAGCTGCAGATTTGATCTGGAGTGGCGCACAGAAAGTGGGGCAAGGA	4687
Db	5177	CCCAACGGGCCAAAGGCTGTGACATGATCTGGAGTGGCGCCAAAGCCGGAATTGCCAAG	5236
QY	4688	TGATCTTGACAGATGAAAGACATCACACCAAGATTTAGATGATTTGAAAGCGACTGAACA	4747
Db	5237	TGGTGTGCAAGACGAAACATTTACACCAAAATAGAGGTGACGTGAAAGCGGCTTTAAACA	5296
QY	4748	CACCTGGCCCATCAACAGGTGCGAATGTTCCGGTGGGCAATTTAGTGTCCAAAGAGGTGA	4807
Db	5297	CACGTATGCAATTAACAGTGTCCAGACGATCCGTGTGGCTCTGTTCTTAAGCAGACCT	5356
QY	4808	CAGCCTATAACGACGTGAACAATTCACCGTCTCCAGACCTCAACAAATTAATATGAAA	4867
Db	5357	CCTCTTAACAATCCCTGCTCTGTGCAAGCATCTTCGACATTCATTGCGAATATGACT	5416
QY	4868	ACATGATCCGGTACACCGGACAGCCCCGACAGCTCCGCTCAAGGACACTTATGATCACTC	4927
Db	5417	CTTCCTTAGGTAACACAGCAGCCCAACACAGCTCCGGTCCCGGTCCCATGATCAACC	5476
QY	4928	CTGACCTTGAAGATGGAAGTCAAGATGTGGCACTTAGTGAAGAACACGACACGAGACC	4987
Db	5477	CAGACTTGGAAGCGGTGTCAAGGTTTGGCACTGTGTAAGAAATATTAACATGATGACC	5536
QY	4988	AGAAGGAGGGGGAACGGGGGAGCAAGATGTGTGAATCTAATCTGACCCGACTCTGG	5047
Db	5537	AGAAGGAGGGTGAACGGGGGAGCAAAATGTGTGTGAATCTAATCTGACCCGGGTCTGAG	5596
QY	5048	CCACTAAGGCACTGCAAGAAATTTGTGATGACTCTTTGAGACCACTTTCAACAGG	5107
Db	5597	CCACCAAGGCACTCTGCAAGAAATTTGTGACGACTTTTGAAGCACTTGTTCACAGACTG	5656
QY	5108	CACACCGTGGCTCTGCTGCCCCCTGGGCATCAATGATTTGATCTCTGTGATGAGC	5167
Db	5657	TGCACCGGGGATGTGCTCTCCCTTAGCAATCAATGATTTGATTTCCGTGATGAGC	5716
QY	5168	AGGCTGATTAACATGGCAATTCATGACCCGACGTCGCCATTACTTGAAGACAAATTGCC	5227
Db	5717	AGGCAAGACGACACGATACCAAGCAACAAGATGTGGGCAACCTGGAAAAAGCAATGCC	5776
QY	5228	TGCCCTGAGGTTTGGGTCAACATGATCAAGAACCCGACAGTTGTGTTGACATCATTA	5287
Db	5777	TTCCACTTCGTTTCTGGGTGAATGTCAATCAAGAACCTTCATATTGTATTGACATGCACA	5838
QY	5288	AGAAAGCATCAACAAGCTGGCTCTCTGTGTGTGGCTCAAGCTCTTACATGACCTTGTCT	5347
Db	5837	AGGGAGATCAACAATGCTGTCTCTCTGTGTGTAGCCCAAGACCTTTATATGACTCTCTTT	5898

OY	5348	CCAGGTGACAGCAACCGGCTGGGCAAGAATCTGCCTTCACAAGACTCTGTATATGCCAAG	5407
Db	5597	CCACATCAGAGCAACCGACTAGGGAAGAACTCACCTTCCACAAAGTGCTGTATGCCAAG	5596
OY	5408	ACATCCCCAGCTACAGAAATTGGGTGGAGAGTTTAATACTACAGATATGGGAAGATGCCAG	5467
Db	5957	ATATCCCAGATTATTAACAATCGGGTAGMAAGAAATACATATGACATATTTGCCAAGTCCAG	6016
OY	5468	CCATCAGCAGCAACAAGATGAAGCATACCGCAATCTGGCTGAGCAAGTCCCGGATCCACATGANT	5527
Db	6017	CCATTATGTACCAAGATATTAATGCTTAATCTCCCGAGCAAGTCCCGCTGCATGCTACAG	6076
OY	5528	AGTTCAACACCATGATGTGCACTCTCAGAGATCTTCTTCCTATGTGGGCAATATACGCGAG	5587
Db	6077	AGTTCAATATGCTGAGGCGCCCTCAACAGAGATCTTAATCTATGTGCAGCAAGATACAGTAGG	6136
OY	5588	AGATCCCTTGACCTCTGACCAACAGATGACCAAGTGTGGAGACGAAMAATCGAGCTACAAC	5647
Db	6137	AGCTATCGGGGCACTTAGCAGAGATGAACAAGGCCGCAACGACATGGCTTACAGG	6196
OY	5648	TAGAACAAAGTCATTAACCTCATGAGCTTTAGACAGCTGAATAAA	5691
Db	6197	TGAGCATCTCATCAACGCGCATGTCCATAGAGAGCTGAAGAA	6240
RESULT 8			
AX704766			
LOCUS	AX704766	6147 bp	DNA linear PAT 04-APR-2003
DEFINITION	Sequence 70 from Patent WO0202634.		
ACCESSION	AX704766		
VERSION	AX704766.1	GI:29561432	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE			
AUTHORS	1		
	Gurujar,R., Hafalia,A.J., Kallick,D.A., Patterson,C., Azimzal,Y.,		
	Khan,F.A., Xu,Y., Yao,M.G., Yue,H., Au-Yang,J., Batera,S.,		
	Baughn,M.R., Borowsky,E.A., Lo,T.P., Lu,D.A., Lu,Y., Tang,Y.T.,		
	Yang,J., Zingler,K.A., Delegane,A.M., Glatzen,K., Marcus,G.A.,		
	Nguyen,D.B., Policky,D.L., Ramkumar,V., Thangavelu,K., Walter,N.K.		
	and Warren,B.A.		
	Human extracellular matrix and cell adhesion polypeptides		
JOURNAL	Patent: WO 0202634-A 70 10-JAN-2002;		
	INCYTE GENOMICS INC. (US)		
FEATURES	Location/Qualifiers		
source	1..6147		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
	/note="Incyte ID No: 7204554CB1"		
ORIGIN			
	Query Match	46.7%; Score 2659.4; DB 6; Length 6147;	
	Best Local Similarity	68.4%; Pred. No. 0; Mismatches 1716; Indels 42; Gaps 8;	
	Matches 3811; Conservative	0;	
OY	141	CGCCGAGGGTTTCATCACTCGTGTGTGATGAGAGACAGACACATTTACTTTGGGCGC	200
Db	302	CGACTGGGGCTCACCCACCTAGTGTCATGAGCAACAAGCGAGGTGATGTGGCGC	361
OY	201	CGTAAATCGATTATTAACGCTTCCACAGACCTGAAGTCTTGGTGAACGATGAGACAG	260
Db	362	AGTAAACCGACTTATATAGCTGTGGGAAACTTATACACTGTGGGGCCACATGACGG	421
OY	261	GCCGACGAGACCAACCCAAAGTGTTCACCAACCCCGCATGCTCCAGACTTGCAATGAGCC	320
Db	422	CCTGTGTGAGAGCAACAAGAAAGTCTACCGCGCCAGCGTGTGAGTCTGCCCCCAGG	481
OY	321	CCTGACCAACCAACATGTCAACAAGATGCTCTCATATGACTTACAAAGAGAACAGGCT	380

Db	482	CTGTGGGAGTACTGACAAACGTCACAAACAGCTGTGTGTGGACATATGCGGTAAACCGCT	541
Qy	381	GATTGCTGTGGAGGCTCTGTACCAAGGATCTGCAGAGTGTGAGGCTGGAGACCTCTT	440
Db	542	GCTGGCCCTGTGTGGACGCGCTCCACAGGGGATGTGGCCAGTTCCTGTGTGGACGATCTCTT	601
Qy	441	CAAGCTGGGGAGGCTTATCATTAAGAAAGACACATATGTCCAGGTGCACAGAGCGG	500
Db	602	CAAACTGGGTATGCCACACACACCGTAAGAGACATACCTGTTCAGCGTCAGAGGAGCG	661
Qy	501	CTCAGCTTTTGAATGATCTCTAC-----AGCAACTGGATGCAAACTGTTTAT	554
Db	662	CAGATGGCGGGCGGCTCATTTGGCGGGCCACCGGGCCAGGGCCAGGCCAAAGCTCTTCGT	721
Qy	555	TGCCACGGCAATGATGGAAGAGCCGGAATATTTTCCGACATTCACACGGGAAACTGAC	614
Db	722	GGGACACCCATCGATGGCAAGTCCGAATCTTCCCACTGTCCAGCCGTGGCTCAT	781
Qy	615	CAAGAACTCTGAGCGGATGGCATGTGGCGTACGTCCTTCATGATGATTCGTGGCCCTC	674
Db	782	GGCCACAGAGGAGATGCGGACATGTGTGGCTTGATACAGATGATTTGTGTATC	841
Qy	675	GATGATTAAATTCCTTTGGACACTTACCATCATCTCTGACTTTGATATCTACTATGT	734
Db	842	ACACCTCAAGATTCCTTTGGACAGCTGTCCAAAGTTCGGGACCTTTGACATCTACTAGT	901
Qy	735	CTATGATTTTASGAGTGGCACTTTGTCTACTTTTGGACCTTCACCTGAGATGATGTTC	794
Db	902	GTAAGCTTCGGGACGAGCAGTTTGTCTACTACTCAGCTGACGCTAGACACAGACT	961
Qy	795	TCCACCAAGCTTCACCCACCAAGAGACAGGTATATCATCCAAAGCTCGTAGAGCTTTTGCA	854
Db	962	GACCTGCGCTGATGCGCGCGGCGAGCACTTCTTACGCTCCAAAGATCGGCGGCTGTGTGT	1021
Qy	855	GGAGGACACAGCCCTTCAATCCATATGTAGAGTGCCATTTGGCTGTAGCGGCAAGGGGCT	914
Db	1022	GGAGGACCCCAATTTCTACTGTACGTTGAAGTTCCCAATTGGCTGTGACAGCAGGCGGGT	1081
Qy	915	GGAGTACCGCTGCTGACGAGCTGCTACTCTGTCCAAAGCGGGGCGGCGTGTGACAGAC	974
Db	1082	GGAGTACCGCTGCTGCTGAGATGCTTACTGAGCGGCGCGGCGGTGCTTGGCCACACA	1141
Qy	975	CCTTGGAGTCCATCCAGATGATGACCTCTCTTCAACCGTCTTTTCCAGGGCCGAAAGCG	1034
Db	1142	GCTGGGCTGTGCTAGAGCAGAGAGTGCTGTTCATCTGTGTTGCGCCAGGGCCGAAAGAA	1201
Qy	1035	GAAATGAATATCCCTGATGATGAGTCGGCCCTGTGCATCTTTCATCTTGAAGCAGATAAATGA	1094
Db	1202	CCGGCTGAAGCCACCAAGAGGTAGCACTGTGCTGTTCACGCTCAGAGGCCATCAGAGA	1261
Qy	1095	CCGCATTAAAGAGCGGCTGCAGTCTTGTTCACCGGGGCGAGGGCAAGCTGACCTTGGCCTG	1154
Db	1262	GAAAGTTTAAAGAGCCCATCCAGTCTCTGCTACCGTGTAGGGGCAAGCTCTCTGCGCTGTG	1321
Qy	1155	GCTCAAGGTGAAGGACATCCCTGTGAGAGTGGCTCTTAACATATGACATTAATCTTGTG	1214
Db	1322	GCTGCTCAACAAGAGCTGGGCTGTGCATACCTGCTCTGCAGATGATGACATTTCTG	1381
Qy	1215	TGGCTTGACATGAATGCTCTCCCTTGGAGTGTCCGACATGATGCTGTGAATTCCTGCTT	1274
Db	1382	CGGGCAGGACTTCAACAGAGCCCTGTGGGGGCGCAGATGCACATTAAGGGGAGCGCCCTGTT	1441
Qy	1275	CACGAGAGACAGGACCGCATGACGTCTGTCACTGCATATGTCTTACAGAAACCACTCTCT	1334
Db	1442	CGTGACAAAGATGATGCTGTGACCGCGCTGTGCTCTATGACTATCGGGGCGGACGTGT	1501
Qy	1335	GGCCTTTTGGGACCAAAAGTGGCAACTGAAGAGATCCGGGTGATGAGCAACA-----	1390
Db	1502	GGTATTCGCCGCGACAGCAAGTGGCGGATCCGCAAGATTCCTGGTGGACCTTCAAAACC	1561
Qy	1391	-----GGGGCAACGCCCTTCAGTATGAGACGGTGACGGTGTGTGACCCCGGCCACAGTCTT	1445
Db	1562	CGATGGCGGCGCTGCTGCGCTTACGAAAGGCTGTGGCCACAGAGGGGCAACCCCATCTT	1621

[illegible]

QY	2523	GGAGAGCAGTGGCTGAGAGCTGTCTGGTGGCCAAAGAAAGAGGCAACAAACCCCGCATCAC	2582
Db	2702	ACCTGCATCGGTGATGACCGCGGCTCACGGGAGCAGTGGCTGACACCGAATCCCAAGATCT	2761
QY	2583	AGAGATAATCCCGGTGACAGGCCCCCGGGAAGGGGGACCAAGGTCACTATCCGAGGGGA	2642
Db	2762	CAACCTGTCCCCGAGACGGGCGCGAGGCAAGGGCGGACCGCGCTCACTATCACAGCGCA	2821
QY	2643	GAACCTGGGGCTGGAAATTTGGCGCATCGGCTCCCATGTCAAGGTCTGTGGCGTGAAGTGT	2702
Db	2822	GAACCTGGGGCTGGAAATTCGAAGACGTGCTGTGGCGGTGGCGGTGGGCAAGGTGCTGTG	2881
QY	2703	CAGCCCTTTAATGATGGTTACATCCCTGCAGAAACAATCTGTGTAGATGGGGGAGGC	2762
Db	2882	CAGCCCTTGGAGAGCGAGTACATCAGTGGCGAGACAGATCCTGTGTAGATCGGGGAGCG	2941
QY	2763	CAAGCCC---AGCCAGCATGACAGGCTTGTGAGATCTGGGTGGCTGTGTGGCTCTGA	2819
Db	2942	CAGCTCTGGGTGGCTCCATGACGCCCTGTGTGAAGTGTGTGGGGAATCTCACACA	3001
QY	2820	ATTGATGGCCCGGTCTCTACAGCTCTATTACTATGACACTGACTCTTCAGATCTGAA	2879
Db	3002	CTACCGCGCCCTGTACCCCAAGCCCTTCACTTGTGACACCAACCTTTACCTGTGTAG	3061
QY	2880	GCCGACCGGGGGGCCCATGTCCGAGAGGAGCCCAAGTACATCAAGGACCAACCTGAA	2939
Db	3062	CCCTTCCCTGTGGCTCTGTCAAGGGGACCTGTGATTTGGCATTCAGAGGAACCACTGAA	3121
QY	2940	TGCCGGAAGCAACGTGTGTGTATGTTTGGAAACAGCCCTGTCTTTCCACAGGCGATC	2999
Db	3122	CGCAGGCGATGATGGCTGTGTGTGTGGTGGCGGCGCCCTGCTCTTCTCTGGAAGAA	3181
QY	3000	TCCATCTCAATGTGCTGCAACACACACATCTCAGATAGGTGTAGAGATGAAGTGTGTC	3059
Db	3182	CTCCCGTGAATTCGGGTGCTGTGACACCCCCGGGCGAGGCCCTGGACGCTCCCATCAT	3241
QY	3060	GGTGAAGGTGACAGGGCCAAAG---ATTCACCAAGACTGTCTTTCAATATGTGGAAGA	3116
Db	3242	CATCAACATCAACCGCGCCCAAGCTCACCAACCTGAGGTGAATGACATCAACCCGAGGA	3301
QY	3117	CCCCCAGCTGTCGGGGAATTGAGCCAGATGAGCAATTGTCAGTGGAACAACCCATCTGC	3176
Db	3302	CCCCACCATCTGTAGAGATCGAACCCGATGAGACATCAACGCGTGGGACCTTCCTGAC	3361
QY	3177	CGTATGGGGAACCACTTGAACTCTATACAGAACCCCAATCCGTGTCAGATGGAAG	3236
Db	3362	GGTCAAGAGCAACCACTGGGCCATGTCTGTGAACCCCGAATCCGGGCAAGTATGAGG	3421
QY	3237	GAAGAGACATCAATATCTGTGAGGTTCTGAACCTACTAGATGACCTGTCAAGCGCC	3296
Db	3422	CATTGAGAGGAGAACGGGTGCTGTGTATCAATGACCAACCATGTATGCGGGCCC	3481
QY	3297	CGCCCTGGCTCTGGGTCTGTACACACAGTCAAGACTGACCGAGAGGCCGAGAGTTTGG	3356
Db	3482	GTCTGTGGCAACCTGTGGGAGCCACCAAGAGTGGGGAGCGGCGGATGAGCTGGG	3541
QY	3357	CTTCATCTGTGACACGTCAAGTCCCTCTCTATCCTCAACAAGACCACTTCACTACTA	3416
Db	3542	CTTGGTATGGAACAAAGTGGGCTGCTGTGTGTGTAATTCACCTCTTCTCTACTA	3601
QY	3417	TCCCAACCCGGTGTGTAGGCTTTGTGTCTCAGGAATCCTGAGCTCAGAGCTGTGAC	3476
Db	3602	CCCTGACCCCGTACGTGAGGACCACTGACCCCACTGGCTGTGTGAGCTGAAGCCGATC	3661
QY	3477	GCCCATCACTCTAAAGGGCAAGAACTATCCCGCTGTGGCTGGGGCAACGTAAAGCT	3536
Db	3662	CCCACTCACTCTCAAGGGCCGGAACTCTTGACACT---GCACCCGGCAACTCCGACT	3718
QY	3537	GAACTCACTGTGCTGTGTGGGAGAAACCGTGAACCGTGAACGTGTCAAGTCTCAAGCT	3596
Db	3719	CAACTCAACGTGTCTCATGTGGTCCACACCTGTATCCTCAACCGTGTGGAGAGCGCACT	3778
QY	3597	GCTCTGCGAGTCCCCCAACTTCATCGGACGACCAAGTATGAGCCGATGTGGTGGCAT	3656

Db	3779	GCCTGTGGAGCGCCGCACTCTCACTGGGAGCAAGGTACAGGTGTGGGCAAGTGGCTT	3838
Qy	3657	GGAGTACTCCCCGGGGATGGTGTACATTGCCCCGAGACGCCGCTCAGCTTGCCGCTCAT	3716
Db	3839	CGAGTTTCTCCACAGGACCACTGACAGGTGTACTGTGGAAGGCTTCTGACGCTTGCTGCAT	3898
Qy	3717	CGTACAGACTGCAGTGGCTGGCGGCTCTCATCATTTTCATGCTGGCCGCTGCTCATATGC	3776
Db	3899	TGTGGCATTTGGCGAAGCGGGGGTCTCTGTCGTGTATCATGTGGCTGTGCTCATCGC	3958
Qy	3777	CTATTAACGCAATGCTCCCGCAAAAGTACCTCAGCGTGAACCGGCTGAGATGAGATGGA	3836
Db	3959	CTAACACCGCAAGTCAAGAAATGTGACCGCACCTCAACCGGCTGAGCTCCAGATGGA	4018
Qy	3837	CAACCTGAGTCCCGTGTGGCCCTTGAGTCAAGAAAGCTTTGCCAGCTGAGACGA	3896
Db	4019	CAACTGGAGTCCCGGTGGCCCTCGAATCAAGAAAGCTTTGCAAGCTGACAGAGA	4078
Qy	3897	CATCCAGAGCTGACCACTGACTGTGATGAGACCGGATTCGCTTCTGTGACTATAAAC	3956
Db	4079	CATCCAGAGCTGACCAATGACTGTGACGTGGCGGATCCCTCTTGTGCTAACCGAC	4138
Qy	3957	TTACACCATGGGGTGTGTGTCCAGGAATTGAAAGCAACCTGTGCTCCGGACCTTGA	4016
Db	4139	ATTAGCATGGGGTGTCTTTCTGGATCGAGAGCAACCTGTGCTCAAGGAGATGGA	4198
Qy	4017	GGTCCCGGGTACCCGGAGAGAGGTGTGGAAAGGCTGAAGTCTTCCCGCAGCTCAT	4076
Db	4199	GGT-----GGAGGCATGTGGAGAAAGTCCCTGACACTGTTCCGGCAGCTGCT	4246
Qy	4077	CACACAAGGATGTTCTGTGCTGTCTTCATTCGCAAGCTTGAGTCCACGCTAGCTTCTC	4136
Db	4247	GACCAAGAAAGCACTTCTGTGACTTCATTCGCAAGCTTGAGGCAACGCGCACTTCTC	4306
Qy	4137	CATCCGGGACCGTGGCAACGTGGCTCACTCATATACCGTGTGACAGCAAGCTGGA	4196
Db	4307	CATCGCGACCGCGGAATGTGGCTGTCTCATATACCGGCTTGACAGGGCGAGATGGA	4366
Qy	4197	GTAAGCCACTGATGTGTGAAAGCAAGCTGTGGCCGACCTCATTTGACAAAGACTGGAG	4256
Db	4367	ATAGGCCACAGGCTGTGCTCAAGCAGCTGTCTTCCGACTCATCGAAGAAAGCTGGAG	4426
Qy	4257	CAACAACCCCTTAGCGTGTGCTCAGAGGAGCTGACGTGGCTGAGAAAGTGTGAC	4316
Db	4427	CAGAAGCACCCCAAGCTGTACTGTCCGCGGAGCTGAATCGGTGGCAGAGAAAGTGTAC	4486
Qy	4317	CAATTGGTTATCTTCTCTCTCTTACAAGTTCTTCAAGAGTGTGTGGGAGCCCTCTT	4376
Db	4487	TATCGGTTCACCTTCTCTTGTATAGTTCTTCAAGAGTGTGCGTGGGAGCGCGTGT	4546
Qy	4377	CTCCGTGTTGCGTGCATCAAGCAGAGATGGAAGAGGCGCCATTGACGCATCAGCGG	4436
Db	4587	CATGCTGTACTGCGCCATCAAGCAGAGATGGAAGAGGCGCCATTGACGCATCAGCGG	4606
Qy	4437	CGAGGCCGCTACTCTCTTGAGCGAGGACAAAGTCAATCCGACGAGATTGACTCAAAAC	4486
Db	4607	TGAGGACGCTACTCTCTGTAGTGTGAGCAAGCTCATCCGGAGAGATGTACTCAAGAC	4666
Qy	4497	CCGTGATCTGAGCTGTGTCAAGCCAGACCAATGCAACAGCCCGAGATCCCAATAAGAT	4556
Db	4667	ACTGACCTGTACTGTGTGAACCTGTGAATGAAATGCAAGCTGAGGTGCGGTGAAGGG	4726
Qy	4557	CCTCACTGTGACCACTCATCAGGTGTCAAGAGAAATTTGATGTCCATCTTCAAGAA	4616
Db	4727	GCTGACCTGTACCGGTCAACCGGCCAAGAGAAAGCTGTGAGCGTGTCTTCAAGGG	4786
Qy	4617	TGTGCTTGTCTCCACCGGCCCAAGAGCTGAGATATGATCTGAGTGTGCGACAAGAG	4676
Db	4787	CGTGCGCTACTCCAGCGGCCAAGCGCGGAGCATGGACTGTGAGTGTGCGCCAGGGCG	4846
Qy	4677	TGGGCAAGATGATCTTGGAGATGGAAGCATCAGCACCAAGATTGAAATGATTTGGA	4736

[illegible]

KEYWORDS	plexin 1.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Kameyama, T., Murakami, Y., Suto, F., Kawakami, A., Takagi, S., Hirata, T. and Fujisawa, H. Identification of a neuronal cell surface molecule, plexin, in mice Biochem. Biophys. Res. Commun. 226 (2), 524-529 (1996)
AUTHORS	1 (site)
TITLE	2 (sites)
JOURNAL	Shimizu, M., Murakami, Y., Suto, F. and Fujisawa, H. Determination of cell adhesion sites of neuropilin-1 J. Cell Biol. 148 (6), 1283-1293 (2000)
MEDLINE	20191894
PUBMED	10725340
REFERENCE	3 (bases 1 to 6466)
AUTHORS	Fujisawa, H.
TITLE	Direct Submission
JOURNAL	Submitted (07-AUG-1996) Hajime Fujisawa, Nagoya University Graduate School of Science, Division of Biological Science, Furo-cho, Chikasa-Ku, Nagoya, Aichi 464-01, Japan (E-mail: fujisawa@bio.nagoya-u.ac.jp, Tel: 052-789-2978, Fax: 052-789-2979)
FEATURES	Location/Qualifiers

CDS

ORIGIN

```

/codon_start=1
/product="p1exin 1"
/protein_len="8413188 1"
/db_xref="ig1.1665757"

/tranlation="MPLPLSPSTYLLSLTLLRLVWYLAISPSPADNGKCPPEAFPTFA
SDMGWTHLVNHEQIQEYVYGAANVRITKLISGNLTIRAWYTPGDEINDEKCTPPQSVG
SHGIGSTNDVNLKLLDLYDAANRLILAKGSSVSGI CQFLDLDDFLKIGEHKHEKYLSS
VREGSSNAQYLIPSPQGOQALFVTPITPDESEYFPLTSSRLMANEEDADMGFYVY
ODEVSSQQLIPSTLTSKSPAFPIYVYVFSBQEQEYVYLTQDLPDTSPLDAHEGVEY
TSKIVLRACNDKPKYSPVSEALICEDQAEYKIDODALYSPPQALAKQGLADEEYFV
LFTFAPQGOKNRKYPPRESALCFLRAIKETIKRISQSGEKSLLPMLLENKQV
CINSPLOIDDPCGQDPNQLAGTAYITETPLFYDKEGILANVAYVQEQYVYFVAT
RSGKIRKYLVDLAPSGRPMALYSVYAEQNPILDLVLSBNQIYLAINTERQYVQ
FVSECVQYTSCELSRDPHCGWCVLHSTGRDQCAERAEPPQFADLIQCVQVTF
QPRNVYSVYMSQVPLVLOANWVDPILKSGVSPENDEETTESILIEBGRHCPMSKRYH
PITQGOQDRVYVLYLAKSTKGTQKSAVDFVFNYSVQSICLALNCGFPHQCKRYH
VCTNNADCAFLFESRWMSDCCQLSPSTHIVPVGVKPLTAAINLPQCSQSGQRYD
NLOAHYIKCALSQSCGLCKLPDPRGECQAEERCKLRHHCNDAPSAMSEAHHS
SRCTDPKILKSPETGROGRTLTITGNEALREBDYRLCHYKQDSCSYSEYYS
AEQVCEIGASLTPRADALVEYCAVDCSLHRAISPPKRFVLTPIYRSPSGSLS
AGTWTGIEKSHLNGSVYISIGCRSPMSWNSIEICLPHPGPESATVITINIRK
AOLNPEKKNYVEDPILIRIDEMRSINGGTLITVYTNALVAPAEPIRAKQIGYD
ENSQVYNDITWVCAEASIDNPKRSPGEPBEGITLIMVNRVLTANSSSLTEYLP
DPVLEPISPLGILIEKRSPLIKGNLLPAPKNSPGLMYLTISCTPLTSEVETOL
LCEAPNLTQGHQYKLVAGGEPSPGLQVYSSNLTPLAYVIGIGGGGLILVAVIPL
TAYRKSKDAPRTKRLQLODMNEVLSRVALECKAPABLIQDTHETLSLDDGATPIL
YIYRYARVYVPEGIEHPVLKEMEVANVEKSLTIFGOLLTKGHLTFTITTELAGQSS
FMSMDRGVNSLIMTALQOMEVYATVLQQLSDLIEKRLSKNAPKILERTSVALE
KQMLTNWFLPILYFKLEKQAGEPFLMYLAKIKQMKGPIDITGEARVLSDEKLIKQ
QIDKXTLLTANVNEEDNAEYVGNKCNCDITVQYKELDLAVYGVSPQSPRAGMDQ
DLERQOGMARILIHEDDVTYKIDNMKCNLTILAHQVYDGSVALEKQASINAVI
SFTFKLIRSEWELTPRASPSPRETPMITPDESGKRLMHLKXNDHLDDEQGR
GSKRWSELVYLRLLATKGTOKVDDLPFTITSLHRSGLALPLAKWMDLDOAGR
HOIHSDYRHRHYSNCLPLAFRYVYIKQFQVPIHNSITDACLISVAQTFMDSCT
SEHKQGOBSPENKLYAKDIPKYSWSEYVADIAKAPISQDMSATLQASJLHLS
PENSMSALHYSYIACYACQDILVALAEKDBQARRRLRSKLEQVNDTALS"

```


Query Match 45.3%; Score 2578.4; DB 10; Length 6466;
Best Local Similarity 67.5%; Pred. No. 0;
Matches 3760; Conservative 0; Mismatches 1766; Indels 42; Gaps 8;

QY 142 GCGAGGGTTTCATACCTGTGTGTGATGAGAGACAGACATTTACTTGGGGCC 201
DB GACTGGGGCTTCACTCACTGGTGTTCACGAAACAGACAGAGATATATGTGGCCT 459
QY 202 GTCAATCGATTTTACAGCTTCCAGCCAGCTGAAAGTCTTGTATACGATAGACAAG 261
DB GTGAACCGTATTTTACAGCTGTGGGAACTTACCTGCTTGGGGCCATGTGACAGG 519
QY 262 CCGAGACGAGACACCCCAAGTGTTCACCAACCCGATGCTGCAGACCTGCATAGACC 321
DB CCGTGAAGACATAGAAATGTCTACCAACCCCGAGTACAGTCCGGCCACATAGT 579
QY 322 CTGACCAACCAACATGTCAACAAAGATGCTCTCATAGACTTACAGAGAAACAGCTG 381
DB CTGGGACAGACAGACAAAGTCAACAAAGCTTCTTACCTGTGACTATGACCCACGCTG 639
QY 382 ATTGCTGTGGAGCTGTACCAAGCATGTGCAAGCTGCTGAGGCTGAGGACCTTTC 441
DB TTGGGCTGTGAGAGTCCCTCAAGGACATCTGCACTTCTGAGTGGACATCTCTTC 699
QY 442 AAGCTGGGGAGCTTTATCATTAAGAGACATATCTGTCAAGTGTCAACGAGCGGC 501
DB AAGCTGGGGAGCGGACACACCGCAAGGAACTACCTGTCCAGTGTTCAGAGGCTGGC 759
QY 502 TCAGTCTTGGAGTATCGTCTCTA-----CAGCAACCTGGATGACAAAGCTGTCT 555
DB AGATGCGGGAGTGTCTATGTGTGGGCAACCGGCAAGGCAAGGCAAGCTCTTCGT 819
QY 556 GCCACGGCAGTGTAGTGGAAAGCCGAGTATTTTCCACCATCTCAAGCCGAAACTGACC 615
DB GGCACACCCATGATGTGCAATACAGAACTCTCCCACTGTCCAGTCCGCGCTCAAG 879
QY 616 AAGAACTGTAGGCGGATGAGCATGTTGCGGTACGTCTTCATGATGAGTGTGGCCTG 675
DB GCCACGAGAGAAATGTCTACATGTTTGGCTTTGTGTACAGAGATGATGTGTCTCAT 939
QY 676 ATGATTAAGATCCCTTGGACACTTACCATCATCTCTGACTTGTATATCTATATGTC 735
DB CAGCTCAAGATCCCTTCAACACATTTGTCAAGTTCACGACCTTGTGACATCTATATG 999
QY 736 TATGTTTACAGATGAGCACTTGTCTCTTTTGAACCTTCAACCTGAGATGTGTCT 795
DB TACAGTTTCCGACGAGCATGTTCTGTATCTACTCACTGACCTTGAACCTTGAACGAGCTT 1059
QY 796 CCACAGGCTTCAACACCAAGAGCAGTGTATATCAATCCAAAGCTGTGAGGCTTTCAG 855
DB ACCTGCCCGACGCGCTGGGAGACCTTCACTCACTCAAGATGTGGGCTGTGGCTG 1119
QY 856 GAGAGACAGCCTTCACTCTCTATGTAGAGTGGCCATTTGGCTGTGAGCGCAGTGGGCTG 915
DB AATGACCCCAAGTCTTACTTCTTACGTGAGTTCCTCCATTTGGCTGTGAGCGCTGGCTG 1179
QY 916 GAGTACCGCTGTGAGGCTGCTACTCTGTCCAAAGCGGGGCGCTGTGGCAGGACC 975
DB GAGTATCGGCTGTGAGGACGCTTACCTGAGCCGCGGAGCAGGCGCTTGGCCAAAG 1239
QY 976 CTTGAAGTCCATCCAGATATGACTGCTCTTCAACCGTCTTCCAAAGGCGCAAGAGCG 1035
DB CTGGGCTGTGAGGATAGAGGCTGTGTTCATCTGTGTTCGACAGGCGCAAGAAAG 1299
QY 1036 AAAATGAATCCCTGATAGTGGGCTGTGATCTTCACTTCTTGAAGAGATATATAC 1095
DB CCGGTGAAGCCACCAAGAGTGGGCTGTGCTGCTGTTCACACTGAGGCGCATCAAGAG 1359
QY 1096 CGCATTAAGAGGCGCTGCAAGTCTTGTACCGGGGCGAGGCGACCTGAGCTGGCTGG 1155
DB AAAATCAAGAGGCGCATCAATCTGTCTACCGTGAAGAAAGAACTCTCTGCTGGCTGG 1419
QY 1156 CTCAGGTGAAGACATCCCTGACAGAGTGGCTCTTACCATTGAAGATATCTTGT 1215

DB 1420 CTGCTCAACAAAGAGTGGCTGTGATCACTCCGCTTGAAGATGTATGACATCTTTC 1479
QY 1216 GGCCTGACATGAAATGCTCCCTTGGAGTGTCCGACATGCTGGGTGGAAATTCGCTTC 1275
DB GGGCAGGATTTCAACGACCATGGGTGGACCGGTCAACATGAGGGACACCCCTTTT 1539
QY 1276 ACGAGGACAGGAGCCGCAATGAGTGTGTATGCAATGTCTTCAAGAAACCATCTCTG 1335
DB GTGCAAGAGAAATGAGCTTGAATGCTGTGGCTGTGCTCTATGATCAAGAGCGACCTG 1599
QY 1336 GCTTTTGTGGCAACAAAGTGGCAAGCTGAAAGATCCGGGTGATGAGACCCAGGGCC 1395
DB GTTTTGTGTGCACTGACAGCGCCGATCCGCAAGTTCGTGTGACCTTGCACAAACCC 1659
QY 1396 AAC-----GCCCTCCAGTATGAGCGGTGACAGTGTGAGACCCCGGCTCAAGTCTC 1446
DB AGCGGACGCGCGCTGTGATGAAAGTGTGTGCTTCAAGAGGCAACCTTATCTG 1719
QY 1447 CCGGATATGCTCTTCCAGAGACACAGCACTCTATCATGTCTAGAGAGCAGCTC 1506
DB CTGACCTGTCTCAGACCCCAACCGCATCTCTATGCTATGAGAGAGAGAGTGT 1779
QY 1507 ACCAAGTCCCTGTGAGTCTGTGTCACTATCAAGTGTGCGGCGAGTGTCTGCTCA 1566
DB ACCGAGGTGCTGTGAAAGCTGTCTCAAGTATCTTCTGTGAGCTGTGCGGCTCT 1839
QY 1567 GCGACCCCACTGTGAGTGTGTGTGTGCAACACAGTGCACCCGAAAGAGCGAGT 1626
DB CCGGACCCCACTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1899
QY 1627 GAGCGGTCAAGAGCCCGCAGCTTGTGCTGGAATGAGAGTGTGTGTGCTGAGC 1686
DB GAGCGGGCAGAAAGCCCAAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1959
QY 1687 GTCCATCCCAATATCTCCGTCTCAGTATCAAGTGTGTGTGTGTGTGTGTGTGTGT 1746
DB GTCCACCGGCAATGT 2019
QY 1747 AATGTCCGAGTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1806
DB AATGTGTCTACCTTCACTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2079
QY 1807 GGGCTGTGTGGCAATCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1866
DB AGCATCTGTGAAGATGGCGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2139
QY 1867 ATCATCAAGAGATGGGACCAACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1926
DB ATCAACAGGGTCAGGGAAGCAGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2199
QY 1927 GGCATGACCTTGGCAGACACAGCTTGTCTTCAATTTGACGAGTGTGTGTGTGTGTGT 1986
DB GGAAGAGT 2259
QY 1987 CTGTCTGTGTGAGATGTCATACCGCTGTCACTGTGTGTGTGTGTGTGTGTGTGTGT 2046
DB CTGGCATGTGTCAACGAGCTCTTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2319
QY 2047 CATGACCCCAAGACTGTCTTCTTCCAGAAAGCGGATGAGTGTGTGTGTGTGTGTGT 2106
DB AACAATGCGGTGATGCTGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2379
QY 2107 CAGCTGTGTGAGTGTGACAGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2166
DB CAGATCTGTCTTCTTCAACCAATCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2439
QY 2167 AAGGCCAAGAACTTCCCGCAGCTGTGGCAGCTGTGTGTGTGTGTGTGTGTGTGTGT 2226
DB GTTCAACGAAACTTGTGCGACCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2499
QY 2227 ATTCAAGGACAGAGAGAGT 2286

D	2500	ATCCAGGCAAGCCGGGCCCGGAGCTACTGCACTGGCGCTTCAACAGCTCCAGCTGAGTGC	2555
Q	2287	CAGAACACCTCTTATTTCTTATGAAAGGATGAGATCAACACTGCCCTGAGATTGCA	2348
D	2560	CAGAACTCTCTGTACTCTTATGAGGGGATGATGTGCACGATCTGCGCTGTAACTGTGC	2619
Q	2347	GTCGTGGAATGAGCACTTCAACTTCAACAACCCAGCTCAGATTAAGTTTCACTCTTAC	2406
D	2620	GTTGTGTGAATGGCACTTTGTCTATTGACAAACCCCGAATCATCAAGGCCCACTCTAC	2679
Q	2407	AAGTGTGAGCCATTCCTGTGAGAGCTGCGGGCTGTGCTTCAAGCTGACCCAGATTGCA	2466
D	2660	AAGTCCCAAGCTTTCGCGCAGAGTTGTGACTCTGCTCTCAAGGCAAGACCCAGCTTTGAG	2739
Q	2467	TGTGCTGTGTCAGAGGCGCCAGGCGCAGTGCACCTTGCGCCAGCACTGCC---CTGCGCAG	2523
D	2740	TGCGGCTGTGTGTAGCTGAGGCGCGGTTCCTTCGTCAACCACTGCGCAGCTGACTCA	2799
Q	2524	GAGAACCACTGTCTGTGAGACTGTCTGTGTGCAAAAGCAAGTGCACAAACCCCGCATCA	2583
D	2800	CCAGCTCTCTGTGATCATGTGCCCAACATGACAGACCGCTGTACAGACCCCAAGATTCTC	2859
Q	2584	GAGATPATCCCGGTGACAGCGCCCGCGGAAGGGGCGCAACAAGTCTACTATCCAGGGAG	2643
D	2860	AAGCTGTCCCGCAGACAGCGCCCTTGGCAGGGAGGAGCTGCACATAACCGGTGAG	2919
Q	2664	AACCTGGGCTCTGGAATTTTCGCGACATGCGCTCCCATGTTCACAGTTGCTGGGTGAGTGC	2703
D	2920	AATCTGGGCTCTCGGTTGTGAGGATGTGCGGCTGAGCGTGCAGTGGGCAAGGTGCTTTGC	2979
Q	2704	AGCCCTTTAGTGATGTTTAACTATCCCTGCAGAACAGATCGTGTGAGATGGGGAGGC-	2762
D	2980	AGCCCTGTGAGAGAGGAGTAACTATCAGCGCAGAACAGATTGTCTGGAGATTGGAGATGC	3039
Q	2763	--CAAGCCCAACAGCATGACAGGCTTCTGTGAGATCTGCGGTGTGTGCGCTGAA	2820
D	3040	AGCAGCTGTGCGGCGCCCATGATATGCTCTGTGGAGGTGTGTGCGGAGCTCTCTCCAC	3099
Q	2821	TTCAATGCGCCGATCTCTCAAGCTCTATTACTTATGACATGACTCTTCTCATGTGAG	2880
D	3100	TACGAGCGCTGTCCCTTAAACGCTTCACTTTGTGACACGACCTTCTACGATGAGC	3159
Q	2881	CCGACCGGGGGCCCATGTCCGGAGAGGCCCAAGTGAACATCAAGGACCAACTGAT	2940
D	3160	CCCTCCCTGGGCTCTGTCTGAGAGTATCTGTGATGGCATGAAGGAGCCATGTGAT	3219
Q	2941	GCCGGAACCAAGTGTGTGATGTTTGGAAAGCAGCCCTGTCTCTTCCACAGGCGATCT	3000
D	3220	GCTGGCAATGATGTGTGCTGTGCTTCATTGGCGGTGCGCCCTGCTCTTCTCTGGAGAAAC	3279
Q	3001	CCATCTTCAATTGTCTGCACACCAACATCTCAGATGAGGTGCTAAGATGAAGSTGTGC	3060
D	3280	TGACGAGAGATTCCTGCTGCTGACACCGCCCGGACACTCTCTGGACGCGCCCTTATGTT	3339
Q	3061	GTCAGGTGACAGGCGCCAAAT---CCACAGAGCTGTGCTTTCAATATGTGAAAGC	3117
D	3340	ATCAACATCAATCGTGCCAGCTTCCAAACCCCGGAGTGAAGTCAACTAATACTGAAGAC	3399
Q	3118	CCGACCATGTGCGGATTTAGCAGAAATGAGAGCATGTTCAGTGGAAAACACCCATGGC	3177
D	3400	CTTACTATCTGAGGATGACCCCTTGAATGAGACATCAATAGTGTGGGACCTCTCTAAC	3459
Q	3178	GTAATGGGGACCCACCTGAGCTCATACAGAAACCCCGAGATTCGTGCCAAGCATGAGGG	3237
D	3460	GTCACGGGACCAACCTGGCAACTGTGCTGTAACCTCGAATCCGGGCAAGTATGAGAGC	3519
Q	3238	AAGAGCAATCAATATCTGTGAGGTTTCTGAACGTACTGAGATGACTCTGACAGCGCC	3297
D	3520	ATTGAGAGGAAATATGCTGATGTGTATCAATGACACACATGTGTGTGCGGCGCGC	3579
Q	3288	GCCCTCGCTGTGGGTCTGACCAACGATCAAGTCAAGTCCGAGAGGCCCGAGAGATTGGC	3357
D	3580	TCTATGACAAACCAAGCGAGCCCGCCAGAACTGGGAGAGCGGCCAGATGAATGTGT	3639

QY	3358	TTCAATCCGGAACAAGCTCAATCCCTGCTCATCTCTCAACAAGAACCAATTCACTACAT	3417
Db	3640	TTTATCATGGAACAAATGTCGCAACGCTGCTGTAATCTCAACAGCTCTCTTCTCTACTAC	3689
QY	3418	CCCAACCCGAGTGTGAAGCCTTTGGTCTCTCAGAAATCTCGAGCTCAACCTTCGACG	3477
Db	3700	CCGGACCTCTGACTGAGGCACTTAAGCCCTACTGCGGCTCTTGGAGCTGAAAGCCAGCTCC	3759
QY	3478	CCCATCATCTCAAAAGGGGCAAGACCTGATCCCGCTGAGGTGGGGGCAAGTGAAGCTG	3537
Db	3760	CCGCTCATCTCTCAAGGGGCGGAATCTCTGTCACT---GTTCTTGCAATTCGCCGCTC	3816
QY	3558	AATCACTGTGCTGTGTGGGAGAAAGCCGTGACCGTGACCTGTGATGATGTCCAGCTG	3597
Db	3817	AACTAACGGGTCTCATCTGCTCCACACTTTGATCTCTCACCGGTCTGAAACACAGCTG	3876
QY	3598	CTGTGCGAGTCCCCCAACCTCATGCGGAGGCAAAAGATAGGCGCGGTGGAGGACATG	3657
Db	3877	CTGTGTGAGAGACCCCAACTTCACGGGGGACGACAAATGACAGGTGCGGGCCGGGCTTC	3936
QY	3658	GAGTACTCCCCGGAGATGTGTACATTGCCCGGACAGCCCGCTGACGCTGCCCATC	3717
Db	3937	GAGTTCTACCTGGGAATGCTGACAGATGTATTCGACAGCCCTGTGACACTGCGGCATT	3986
QY	3718	GTCAGCATCGAGTGGGTGGCGGCTCTCTCATTTTCATCTGTGCGCTGCTCATTTGCC	3777
Db	3997	GTGGGTATCGGTGGTGGGGGTGGGCTCTGTGCTGTGTATCTGTGCTGTCTCATGCC	4056
QY	3778	TATTAACGCAAATGCCCGGAAATGACCTCAAGCTTGAAGCGGCTGCAATGAGTGCAC	3837
Db	4057	TACAAACGCAAATCCCGGAGATGTGACCTGTACCTCTCAAGCGGCTGCAAGCTCCACATGAC	4116
QY	3858	AACCTGAGTCCCGTGTGAGCCCTGAGTGCAGAGAAAGCTTGGCGAGTGCACAGGAC	3897
Db	4117	AACCTGAGTCAAGGGTGGCCCTCGATGTCAGAGAACTTTTGAGAGCTACAAACGAC	4176
QY	3898	ATCCATGAGTCAACCATGATGACTTGATGAGCCGGGATTCCTGTTCTGTGACTATGAAT	3957
Db	4177	ATCATGATGCTGACCAAGTACCTGGAATGAGTCTGCACTCCCTTCCTTGACTTACCGACA	4236
QY	3958	TACACATGCGGGTGTCTTCCAGGAATTTAAACACACCTGTCTCTCCGGACCTTGAG	4017
Db	4237	TATGCGATGCAAGTGTCTGTTCTCGGATTTGAGACCATCTGTACTGAAAGATGAG	4286
QY	4018	GTCCCGGCTACCGGCAAGAGCGTGTGAGAAAGCTGAAGCTTTCGCCAGCTATC	4077
Db	4297	GT-----GAGGCGCAATGTGAGAAAGTGTGACACTGTTCGACACTGTGTG	4344
QY	4078	AACAACAAGGTGTCTGCTGCTCTTATCATCGGACGCTGATGCCAGCGTATCTTCTCC	4137
Db	4345	ACCAAGAAAGCATTTCTCTGTTAACCTTATCCGCACATTAAGGCGCCAGCGTATCTTCC	4404
QY	4138	ATGGCGACAGCTGGCAACGTGGCTCTCATCATTAACCTGTCTGACAGACAGACTGAG	4197
Db	4405	ATGGGTGACCGTGGAAACGTGGCTCTCATCATCAATGACAGCCCTGCAAGGTGATGAG	4464
QY	4198	TACCCACTGATGTGTGAAGCAAGCTCTGGCCACTCATTTGACAAGAACTTGAAGC	4257
Db	4465	TATGCGACAGGATGTCAAAAGCAATGTCTGTGTGACCTCATTCAGAAAACCTTAGAAGC	4524
QY	4258	AAGAACAACCTTAAGCTGTGCTCAGAGAGACTGATGATGAGTGTGAGAAATGCTGACC	4317
Db	4525	AAGAACAACCCCAAGCTCTCTGTTAACGAGAGACCAAACTAGTGGAGAGAAATGTTGACC	4584
QY	4318	AATGTGTTACTTCTCTCTCTCAAAATTCCTCAAGAGTGTGCTGGAGAGCCCTCTTC	4377
Db	4585	AACTGTGTACCTTCTCTTTGTATAATTTCTCAAGGATGTGTGAGGAGCGCTCTTT	4644
QY	4378	TTCCTGTTCTGTGCATCAAGACAGATGAGAGAAAGGCCCATTTGACGCTATACGGGC	4437
Db	4645	ATGCTGTACTGTGCATCAAGACAGATGAGAGAAAGGCCCTATTTAGCGCTATACAGGG	4704

QY 4438 GAGGCGGCTACTCTTGAAGGAGACAGCTCATCCGCGAGATTTGACTACAAACC 4497
 Db 4705 GAGGCGGCTACTCTTGAAGGAGACAGCTCATCCGCGAGATTTGACTACAAACC 4764
 QY 4498 CTGGCTCGAGCTGTGTCAGCCCAACATGCCAAGCCCGGAGTCCAGTAAATGC 4557
 Db 4765 CTGACCCCTGAAGTGTGTCAGCCCAACATGCCAAGCCCGGAGTCCAGTAAATGC 4824
 QY 4558 CTCAACTGTGACCATCACTCAGTCAAGGAGAAATTTGATGCCATTTAAAGAT 4617
 Db 4825 CTGAAGTGTGACCATCACTCAGTCAAGGAGAAATTTGATGCCATTTAAAGAT 4884
 QY 4618 GTGCTTGTCTCCACCGGCGCAAGCTCAGATATGATCTGAGTGGCGCAAGAGAT 4677
 Db 4885 GTGCTTGTCTCCACCGGCGCAAGCTCAGATATGATCTGAGTGGCGCGAGGTCG 4944
 QY 4678 GGGGCAAGATGATTTGAGAGATGAGACATCACCAACCAATGATGATGATGAG 4737
 Db 4945 ATGGCAGCATCATCTTCAGGAGACGTCACCACTAAATTTGACATGACTGAG 5004
 QY 4738 CGACTGAACACACTGGCCCACTACAGGTTGCGAGATGTTCCGTGGTGGCATTTAGTCC 4797
 Db 5005 AGGCTGAACACACTGGCTCACTATAGTGAAGTGGTGTATGGTGGCATTTAGTCC 5064
 QY 4798 AAGCAGGTGACAGCTTATAGCAGTGAACAATTCACCGTCTCCAGGACCTGACAGT 4857
 Db 5065 AAGCAGCATCGGCTTACAC--ATCTCAACTCTCTTACCTTACATTAAGTCCCTCAG 5121
 QY 4858 AATATGAAAACATATCCGGTACACGGGAGCCCGGACCGTCCGTCACGGACACT 4917
 Db 5122 AGATTCGAGAGATCTGTGTCAGGACGCTAGCCCGGACGCTACGTCAGAACCC 5181
 QY 4918 ATGATCACTCTTGAAGTGTGAGTCAAGATGTGGCACTAGTGAAGAACCAAGAG 4977
 Db 5181 ATGATCACTCTTGAAGTGTGAGTCAAGATGTGGCACTAGTGAAGAACCAAGAG 5241
 QY 4978 CACGAGACCAAG 5037
 Db 5242 CACCTGAGCAACAGTGAAGGCGACCGTGGAGCAAGAGAGAGAGAGAGAGAGAG 5301
 QY 5038 CGACTCTGGGCACTAAGGCGACACTGAGCAAGATTTGATGATCTTGTGAGACATC 5097
 Db 5302 AGGCTGTGGGCACTAAGGCGACACTGAGCAAGATTTGATGATCTTGTGAGACATC 5361
 QY 5098 TTGAGACAGGCAACCGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5157
 Db 5362 TTGAGACAGGCAACCGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5421
 QY 5158 CTGATGAGCAGGCTGATTAACATGCAATTCATGACCGGCAAGTCCGATTAAGAG 5217
 Db 5422 TTGATGAGCAGGCTGATTAACATGCAATTCATGACCGGCAAGTCCGATTAAGAG 5481
 QY 5218 AGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5277
 Db 5482 AGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5541
 QY 5278 GACATTCACAAAGACATCAAGAGCGCTGCTCTGAGTGGTGGTCAAGCTTCATG 5337
 Db 5542 GACATTCACAAAGACATCAAGAGCGCTGCTCTGAGTGGTGGTCAAGCTTCATG 5601
 QY 5338 GACTCTTGTCTCAGCTCAGAGCAGCGCTGGGCAAGAGTCCCTCCCAAGAGCTGTG 5397
 Db 5602 GACTCTTGTCTCAGCTCAGAGCAGCGCTGGGCAAGAGTCCCTCCCAAGAGCTGTG 5661
 QY 5398 TTGCGCAAGAGATCCCGGCTCAAGAAATTTGGTGGAGAGGATTAATCAAGACATAG 5457
 Db 5662 TTGCGCAAGAGATCCCGGCTCAAGAAATTTGGTGGAGAGGATTAATCAAGACATAG 5721
 QY 5458 AAGATTCACAGCATCAGCAGCAAGACATGAACGATACCTGGTGAAGAGTCCGAGT 5517
 Db 5722 AAGATTCACAGCATCAGCAGCAAGACATGAACGATACCTGGTGAAGAGTCCGAGT 5781
 QY 5518 CACATGAATGATTTCAACACATGATGACCTCTGAGAGATCTTCTCTATGTGGGCAA 5577

Db 5782 CATCTCAGCCAGTTTAAACAGATGAGTGGCTTGCACGAGATCTACTCTTACATGCCCAAG 5841
 QY 5578 TACAGGAGAGATCTTGGACCTTGGACCAAGATGACAGTGTGGAGCAAGAACTG 5637
 Db 5842 TACAGGAGAGATCTTGGTGGCTTGGAGAAAGATGAAACAGGCTCGAAGGAGAGGCTT 5901
 QY 5638 GCTTCAACTGAGACAGTCAATACCTCCTAGACTTGAAGACTGTA 5685
 Db 5902 CGAAGCAACTGAGAGAGTGTGACACACATGCTGAGGAGCTTGA 5949
 RESULT 10
 HSEXGENE
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 H.sapiens
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 6039)
 Maestri, E., Yamaguchi, L., Tamani, F., Cremona, O., Gulsano, M.,
 Longati, P., Bione, S., Neel, B., Toniolo, D., and Comoglio, P.M.
 A novel family of transmembrane proteins with homology to the
 Mel/HGF receptors
 Unpublished
 2 (bases 1 to 6039)
 Toniolo, D.
 Direct Submission
 Submitted (09-JUN-1995) D. Toniolo, Istituto di Genetica
 Biochimica ed Evoluzionistica, CNR, Via Abbiategrasso 207, 27100
 Pavia, ITALY
 This entry extends and replaces X67197.
 COMMENT
 FEATURES
 source
 1..6039
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="X"
 /map="q28"
 /tissue_type="fetal brain, skeletal muscle"
 /clone_id="lambda ZAP, lambda GT10"
 1..6039
 /gene="sex"
 180..188
 /gene="sex"
 /note="kozak consensus"
 185..5800
 /gene="sex"
 /codon_start=1
 /product="SEX protein"
 /protein_id="CA61132.1"
 /db_xref="GI:1200235"
 /db_xref="COA:P51805"
 /db_xref="SWISS-PROT:P51805"
 /translation="MPSVCLLLLLFLAAGALGNRPRAVVVDTTLTHAVRVTE
 VFGAVNRVFXLAPNITERRAVTGVEDNARCYDPSRVAHAHLAPVNNKILLI
 DVARLVVAGSTWQSTCOPLRDLDFKGEFHKKHKLHLSAOGSDSAKATVYOG
 GBEKLFVGVAVDSEKSEFPTLSRKISDESDAMSESVYQDFVSSQIKTISDTLS
 YPALDIYTVGVASFPVFTLQDIDTQTLDTAGSEFTSKIVRMCGDSEFYSYV
 EPIGCSWGEVRLVQSAHLAKPGLLAQALGVADVDVLTFTISQOGKNSAPRQ
 TILCFTLTNINAHIRRIOSCYRGSTLALPMLINKELPCINTPQOINFCGLVNL
 OPIGLHVEIGLPLADSTDMASVAAYVLRQSVYFVIGTSGSLKTVAVDFQDHL
 YETVPVVDGSPILRDLSPDRHRTILSEKVSQVSPVETGROYSCAALGSDGDC
 GNCVLRRCRGACICGASAPRPAHELSKYOVRVNNVTSRQVLTYYTHVVP
 DLSAGVSCAPEAENEAVLPSGELCSPSLOERALTTRGHARTRLRLISMET
 GVAFAGADVFNCSVLOSQSCVSPYCHCKKTRHTLTSRPHSCFVQGRHNSFBG
 CEILPSGDLIPVGMPLTIRAKLDPQSGQKRYCEVAVQOQKQVPAVRNFS
 SVQCNASYSYEDGHEGTELDPSVVDWDFPIDKPRFALLYKMAORPSCGLCLK
 ADRFVNGMCISEHRCQLRTHCPAPKTNMHLISQKTRCSHRIQTQIHPLVBPKEGT

RTTIVGENILLSREVGIRVAVRNCNSIPAEITSAERIVCEMESLVPSPPGVELC
 VDDCADPRTQSEQVYSFVTPFDQVSPBRGASGTLTISGSLDMSVTVVAD
 SSCQFBRDAPAKIVICISPLSTSPSOAPITLIDRANLSPGLATYVADPVTALB
 TWSIINGSTAIIVSGTHLTVQEPVRVAKIETPNQVINDNAMCKAKGFLGR
 POPRAQEHDPFEGFLDHTVQBRNLNSFRYEDPSEFPGSGVDFVGRSVLGR
 KKNKIIPAAAGSSRLNVTYLGOPCSLTVSDTOLCSPGOTGROPVWVGLBEM
 LGTLHISARALLPLPAMWGLAGGCLLILATVAVKATOPDRPLTXKJLOCMN
 LESRYALECKENAFABELOTDINELTNHMEVOJPELDYTYAVRVLPETIEHVKEL
 DTPVNERKALRLEFGULHSRAVLTPIHTLVAQSSFSMRSDPSTVSLTMVALQSRLDY
 ATGLKQOLADLIERNLESKNPKLLRRTESVAKMTLMFTLHFKFLKACGEPL
 FLVYCAIKQOMKGPIDATIGBARYSLSBDKLIROIDYKTLTLMHFCVPEMBSAOVY
 VLVNDSITTOAKDLIDTVYKGIPIYSORPKAEDMDLEPMQOPMTRILIOBEDVYTK
 ECDMKLNSLAIHVOYDGLVATLVPKOSAVYMANSPFTSGLSYESLSTLAKTSPS
 LSRAPMTIPDOETGTMHVLVKNHNDHDPBPDGSPMSEIYITLPLAKTLOKQF
 VDDLEPVTSTHRSALPLAKTWMFDELDEADROQISDDEPVRITKSNCLPLAFW
 NYTKQFQVFDIHKNSITDACSIVAQFPMDSCTSERILGSLSPNLLYAKQIPNY
 KSMVERYRIDAKMASISDQMDAYLVQSRHLHADFVLSALNELYFVTKYRQELI
 TALDRBASCRKHKLRQKLEQIISLVSSDS"

ORIGIN

Query Match 44.6%; Score 2535.4; DB 9; Length 6039;
 Best Local Similarity 67.3%; Pred. No. 0;
 Matches 3730; Conservative 0; Mismatches 1777; Indels 39; Gaps 9;

QY 152 TCAATCAGCTGTGTGATGAGAGACAGACATTACTTGGGGCCGTCAATCCGA 211
 Db 282 TTACCACCTGGCTGTGACCGGGTGAAGTGTGTGGCCGAGTGAACCGAG 341
 QY 212 TTACAAGCTCCGACCGACCTGGAAGTCTTGTAACGATGACAGGCGGACGAG 271
 Db 342 TCTTTAAGTGGCCCCCACTGACTGAGCTGGGGCCCATATGTCAGGAGCCGCTGCGAG 401
 QY 272 ACAACCCCAAGTGTATACCAACCCCGCATGTCACAGCTGCAATGAGCCCTGACCA 331
 Db 402 ACAAGCTGTGTGATACCGGCCCCAGATGCGGTGTGTGCCACCGCTGTGGCCCCG 461
 QY 332 CCAACATGTCAACAAGATGCTCTCATAGACTACAGAGAAACAGGCTGATGCTGTG 391
 Db 462 TGGACAACTCAACAAGCTGCTCATAGACTATAGCGGCCCGCGCTGTGTGCTGCG 521
 QY 392 GGAGCCTGTACCAAGCATCTCAAGCTGCTGAGGCTGAGAGACCTCTTCAAGCTGGGG 451
 Db 522 GGAGCATCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 581
 QY 452 AGCTTTATATAGAGAGAGACTATCTGTGAGGTGTCAAGAGAGCGGCTCACTTTG 511
 Db 582 AGCCGACACCGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 641
 QY 512 GAGTATGCTGTCTACAGCAACCTGAGTGAACAAGCTGTTCATTTGCCACGCGGAGTGA 571
 Db 642 GTGTCTTTGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 701
 QY 572 GGAAGCCCGAGATTTTCCACCATCTCAGCCGAGAACTGACCAAGAACTGTAGAGCG 631
 Db 702 GCAAGTCGAGATCTCCCACTTGAAGCTGCGGAGACTCATAGTATGAAAGACAG 761
 QY 632 ATGGAGTGTGCGGAGTGTCCATGATGATGATGATGATGATGATGATGATGATGATG 691
 Db 762 CGGAGATGTTCAGTCTGTGTACAGAGATGATGATGATGATGATGATGATGATGATG 821
 QY 692 CGGACACCTTCAACATCTCTGACTTGTATGATATGATGATGATGATGATGATGATG 751
 Db 822 CAGACACGCTGTCTTGTACCTGTGCTTGTACATTAATGATGATGATGATGATGATG 881
 QY 752 GCAATTTGTCTACTTTTGAACCTTCAACCTGAGATGATGATGATGATGATGATGATG 811
 Db 882 CTTCTTGTGTACTTCTGAGCTGTGAGTGAACCCAGACAGCTGTGTGAGACAG 941
 QY 812 CCAAGAGAGGATGATATCAATCAAGCTGTGAGGCTTGTGAGAGAGACAGACTTCA 871
 Db 942 CGGAGAGAGAAATTTTCAAGTCAAGATGTGTGCGATGTGTGCGGAGAGACTCAGAGTTCT 1001
 QY 872 ACTCTATGTAGAGGTGCCATTGGCTGTGAGCGAGTGGGAGTGAAGTACCGGCTGTGCTC 931

Db 1002 ACTCATACGTGGAATTCCTCCATCGGCTGTCTGTGCGCGCGCGCTGTGAGTACCGCTTGCTG 1061
 QY 932 AGGCTGCTTACCTGTCCAAAGCGGAGGCGGTGCTGTGACAGAGACCTTGTGAGTCCATCCAG 991
 Db 1062 AGAGCGCCACCTGTGCCAAGCCTGTGCTGTGTGCGCCAGAGGCTGTGGGTGCTCGGCTG 1121
 QY 992 ATGATGACTGTCTTCAACCGTCTTCTTCCAGAGGCGCAAGACGGAATATGAAATCCCTGTG 1051
 Db 1122 ATGAGAGCGTCTCTTCAACCATCTTCTCTCAGGAGCGCAAGAAACCGGAGCCAGCCACCC 1181
 QY 1052 ATGAGTGGGCGCTGTGATCTTCAATGAGAGATTAATGACCGCATTAAGAGAGCGG 1111
 Db 1182 GCGAGACCATCTCTGTCTCTTCAACCTTCAAGCAATCAATGCGCATTCGCGCGCGCA 1241
 QY 1112 TGAAGTGTGTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1171
 Db 1242 TCAGTGTCTGTATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1301
 QY 1172 TCCCGTGGAGAGTGTGCTCTTAAACATGAGATTAATGATGATGATGATGATGATGATG 1231
 Db 1302 TGCCCTGATCAACACCCCATACAGATCAACGCGCACTGTGTGAGCTGTGTGATTAACC 1361
 QY 1232 CTCCCTGGAGAGTGTCCGACATGATGTGAGATTTCCCGTCTTCAAGAGAGACGAGGAGC 1291
 Db 1362 AGCCTGTGGAGGCTGTGATGTATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1421
 QY 1292 GCATGAGTGTGTATGCAATGTATGTCTAACAAGAACCACTCTCTGTGCTTTGTGGGACCA 1351
 Db 1422 GCATGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1481
 QY 1352 AAAGTGGCAAGCTGAGAGAGAGATCCGGGTGATGAGACCAAGGAGGACAGCCCTTCAATATG 1411
 Db 1482 GCAAGGCGAGCTTGAAGAGGTGCGGGTGTGATGTCTTCAAG---ATGCCACCTGTATG 1538
 QY 1412 AGACGCTGAGGTGTGTGAGACCCCGGCTCACTCTCGGAGATATGCTTCTTCAAGAGAC 1471
 Db 1539 AGACGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1598
 QY 1472 ACAGGCAACTCTCATCATGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1531
 Db 1599 ACCGCAATCTATCTCTGTAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1658
 QY 1532 GTGATATGAGAGCTGTGCGGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1591
 Db 1659 AGCAGTACAGAGAGCTGTGAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1718
 QY 1592 TGCTGCAACAACGTGTGACCCCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1651
 Db 1719 TGCTGTGACACAGGTGTGTGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1778
 QY 1652 TTGCTGTGAGATGAGAGAGTGTGTGCGGCTGACGATTCATCCCAATATCTTCCGTCT 1711
 Db 1779 TTGCTGAGAGGCTGACCAAGTGTGTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGG 1838
 QY 1712 CTGAGTACAAAGT 1771
 Db 1839 GTGTACCTGGGAGTGTGAGCTGTGATCACTGTGACAACTGTGACAGCTTCAAGTGTGAG 1898
 QY 1772 TCAACTGCACTTTGAGACCTGTGTGAGATGAGATGAG---CTGTGTGTGTGTGTGTGTGT 1828
 Db 1899 TGAAGTGTGCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1958
 QY 1829 TCGAGTGTACTTCCCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1888
 Db 1959 TGCTGTGCTGT 2018
 QY 1889 ACCATGTGTAGAGCTTCAAGTCAATCAAGAGAGCCGCGATGACCTTGTGCGCAGACCA 1948
 Db 2019 CCGGACATGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2078
 QY 1949 GCTTTGTCTTCAATTTGAGGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2008

Db 2079 ACTTGTCTCTTCAACACTGACGCTCTCCAGTCGTGCAATGTCCTGTGTTGGACGCCCTT 2138
Qy ACCGTGGCACTGGTGTAAATACCAGGCACTGTGCAACCATGACCCCAAGACCTGCTCT 2068
Db 2139 ACCCTGGCACTGGTGTAAATACCAGGCACTGTGCAACCATGACCCCAAGACCTGCTCT 2198
Qy 2069 TCCAGAAAGCCGAGTGAAGTGTGCCGAGACTGTGCCCACTGTCTGTGCAATGAGACA 2128
Db 2199 TCCAGAAAGCCGAGTGAAGTGTGCCGAGACTGTGCCCACTGTCTGTGCAATGAGACA 2258
Qy 2129 TCCGTGTGCGCGTGGAGGTGATCAAGCTTATCAGCTGAAAGCCCAAGACCTCCCCAGC 2188
Db 2259 TCCGTATCCCGTGTGGGTTCATGACGCTTCTTACCTTGGGGCTTAAGAACTTACTAGC 2218
Qy 2189 CCCAGTCTGGGAGCGGTGCTAAGAAATGATCTCAACATTCAGGAGCCAGAGCTGAG 2248
Db 2319 CGCACTCGGGCCAGAAAGAACTATGATGCTGTGTGCGGGTGCAGAGGCGGAGCGGG 2378
Qy 2249 TGCCCGCCCTGGCTTCAACAGCTCCAGCTGTCAGTGCAGAGCAACCTCTTATCTTATG 2308
Db 2379 TGCCCTGGGTGCTTCAACAGAGCACTGTGCAATGCTGAGAGAGCTCTGTACTCTATG 2438
Qy 2309 AAGGATGGAATCAACAACCTGCCCCGTGAGATTGACAGTGTGTGAATGGCACTTCA 2368
Db 2439 AAGGTGATGAGATGTGACACCGAGCTGGAATCTCCGTGTGTCTGGGATGGAATCTTCC 2498
Qy 2369 ACATTGACAAACCCAGCTCAAGATTAAGTCACTCTCAAGTGTGAGGCCATGGCTGAGA 2428
Db 2499 CCATGACAAAGCTTCCAGCTTCCAGCTTCCCTGTGTACAGTGTGGGCGAGGAGGCCCA 2558
Qy 2429 GCTGGGCGCTGTGCTCAAGCTGACCCAGACTTGTGCAATGTGTGTGTGCGAGGCCAG 2488
Db 2559 GCTGTGGCTCTGTGCTCAAGCTGATATCCCGCTTCACTGTGTGTGTGTGTCAAGAGC 2618
Qy 2489 GCCATGTCACTCTGTGCGCACTGTGCTGCTGCGAGAGAGCACTGTGTGTGTGTGTGTG 2548
Db 2619 ACAGGTGCGAGCTGTGGAGCCCACTGCTGCGGCCCCAGAGCAACCTGTGTGTGTGTGTG 2678
Qy 2549 GTGCCAAAGCAAGTGTGCAACAACCCCGCATCAAGAGATATCCCGGTGTGACAGGCC 2608
Db 2679 AGAAGGCAACCGGTGTGCAACCCCGCATCAAGAGATATCCCGGTGTGACAGGCC 2738
Qy 2609 GGGAAAGGGGCAACAAGTCACTATCCAGAGGAGAGACCTGTGGCTGTGAATTTGGACA 2668
Db 2739 AGAAAGAGGCAACCGGTGTGCAACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2798
Qy 2669 TGCCCTCCATGTCAAGGT 2728
Db 2799 TGCGC-----CTGCGGGT 2852
Qy 2729 CTGCAAGACAGATCGT 2785
Db 2859 GTGTGAGAGATCGT 2912
Qy 2786 TCGTGTGAGATCGT 2845
Db 2913 CCGTGTGAGCTGT 2972
Qy 2846 ATTACTTCAATGACATGACTCTCTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2905
Db 2973 AAGCTTTGTGACCCCAAGTTGACCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3032
Qy 2906 GGAACCAAGTGTGATCAAGAGCAACCACTGTAAAGCCGAAAGCAAGTGTGTGTGTGTGT 2965
Db 3033 GGAACAGGCTTACATGTAGAGAGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3092
Qy 2966 TTGGAAGAGAGCCCTGTCTCTTCAAGAGCAATCTCAATCTTCAATGTGTGTGTGTGT 3035
Db 3093 TTAGGAGACAGAGAGT 3152
Qy 3026 CATCTCTCAATGAGT 3082
Db 3153 CTCTCTTCAACCTGT 3212

Qy 3083 T---CCACGAGACCTGTGTCTTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3139
Db 3213 TCTCCAGCCCCGGGTGTCTTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3272
Qy 3140 CAGAAATGAGATGT 3199
Db 3273 CCACTGTGAGATGT 3332
Qy 3200 TCAATCAAGAACCCGAGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3259
Db 3333 CGGTCCAGAGGCCCGGGT 3392
Qy 3260 AAGTGTGTGAGCTGT 3319
Db 3393 AAGTGTGTGAGCTGT 3452
Qy 3320 ACCAGTCAAGCTGT 3379
Db 3453 AGCTGTGAGGCGAAGCGAGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3512
Qy 3380 CCTGTGTATCTGT 3439
Db 3513 CGGCTGT 3572
Qy 3440 TTGATCTCTCAAGAAATCTGT 3499
Db 3573 TGGGGGCTCTGT 3632
Qy 3500 ACTGTATCTGT 3559
Db 3633 ACTGTAT---TCCGAGGAGCGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3685
Qy 3560 AGAAGCGGT 3619
Db 3690 GCGAGCGGT 3749
Qy 3620 TCGGAGGCAACAAGT 3679
Db 3750 CTGGCGGAGCGCTGT 3809
Qy 3680 ACATGTGCCCCGAGACCGGT 3739
Db 3810 ACATGTGCGAGAGCGGAGCGGT 3869
Qy 3740 GCTTCTCATATTTCATGT 3799
Db 3870 GGTCTGT 3929
Qy 3800 GTGACTCAACGT 3859
Db 3930 CGGACCGT 3989
Qy 3860 TGT 3919
Db 3990 TGT 4049
Qy 3920 TGT 3979
Db 4050 TGT 4109
Qy 3980 CAGAAATGT 4039
Db 4110 CGGAGT 4159
Qy 4040 GT 4099
Db 4160 --GT 4217
Qy 4100 CTTTCAATCGCAAGCTGT 4159
Db 4218 CTTTCAATCGCAAGCTGT 4277

QY	4160	CTCACTCATCATATGACGTCGTGTGAGAGCAAGCTGCGAGTGAAGCCACTGATGTCGTGAAGC	4213
Dp	4278	CTTGCTGCTCAACATGATGTGGCCCTGTGACAGCCGGCTGTGACTGTATGACACGGGCTGTCAAGC	4337
QY	4220	AGCTGCTGGCCGACCTCAATTGACAAAGAACCTGGAGAGCAAGAACCACTTAAGTCGTGTC	4279
Dp	4338	AACTGTGGGGCGAAGCTCTCATGAGAAAGAACCTCGAGAGCAAGAACCACTCCCAAGTGTGTCGC	4397
QY	4280	TCAGAGGACCTGAGTCAATGTGGCTGAGAAAGATGCTGACCAATTGGTTTACTTTCCTCTCT	4339
Dp	4398	TACGACAGACAGAGTCAATGTGGCTGAGAAAGATGCTTACCAATGTGTTACAGTTCCTGTGTC	4457
QY	4340	ACAAATTCTCAAGAGATGTGTGTGGGGAGCCCTCTTCTCTGTCCTGTCGTGTCATCAAGC	4399
Dp	4458	ATTAATTTCTGAAGAGATGTGTGTGGGGAGCCCTCTCTCTGCTTTACTGTGCATCAAGC	4517
QY	4400	AGCAGATGAGAGGGGCCCATTTACCGCATACGGGCGAGGGCCCGCTACTCTTGAAGC	4459
Dp	4518	AGCAGATGAGAGGGGCCCATTTATGTCATCACGGGCGAGGCACGATACTCCTTGAGC	4577
QY	4460	AGGACAACTCATCGCGCAGAGATGTACTCAAAACCTGTGCTGTAGCGTGTGACAGC	4519
Dp	4578	AGGACAACTCATCGTGTAGAGATCGACTCAAGACCTGACCTTTCCTGCTGTGTGTC	4637
QY	4520	CAGACAAATGCCAACGCCCCGAGGTCCTCAGTAAAGATCTCTCACTGTGACACATCACTC	4579
Dp	4638	CGAGAAAGAGGGGAGCGGCCAGGTCCTCAGTAAAGGTTCTCACTGTGTGACAGATCAACC	4697
QY	4560	AGGTCTAGAGAGAAATTTGTGATGTCATCTTCAGAAATGTGCTTGTCTCCACCGGCCCA	4639
Dp	4698	AGGCTAAAGATTAAGCTGTGTGACACTGTGTACAAAGGGCATTCGTATCTCCAGGTGCCA	4757
QY	4640	AAAGCTGCAGATATGGATATGTGATGTGGCCACAGAGAAATGGGGCAAGATGATCTTTGAGG	4699
Dp	4758	AAAGCTGAGGATATGACCTGTGATGTGGCCAGGGCCGAGTATCACTGCATATCTTCAGG	4817
QY	4700	ATGAAGACATCAACCAAGATTGAGATTGTGAGACGACTGACACACTGGCCCACT	4759
Dp	4818	ATGAGGATGATCAACCAAGATCGAGTGTGACTGGMAAGGCTCAATCACTGTGCCCACT	4877
QY	4760	ACAGAGTCCAGATATGTTCCGATGGGGGACATTAGTCCAAACAGGTGACAGCTATACG	4819
Dp	4878	ACAGAGTACAGACGGTTCCTTGTGTGGGATTTGTGTGCCAAACAGTGTGTGCTATACCA	4937
QY	4820	CAGTGAACAATCTCAACGCTTCCAGAGACTGACAAAGTAAATATGAACAATGATCCGCT	4879
Dp	4938	TGGCCAACTCTTCACTTCAACCG---CTCCCTCAGCGGCTGACGAGAGTGTGCTCGCA	4994
QY	4880	ACAAGGGAGAGCCCGACAGCCTCGGCTACAGGACACCTATATATCACTCTGACCTGAGA	4939
Dp	4995	CGGCTGAGAGCTGTATAGCTTCGCTCACGGGCAACCATATATTCGCTTACCAAGAGGA	5054
QY	4940	GTGAGTCAAAATGTGTGCACTTAGTGAAGAACCAAGACAGGAGACCAAGAGAGGGGG	4999
Dp	5055	CAGGACCAAAATGTGTGGCACTGTGTGAAGAACCAAGACCATGTGGACCATGTGAGGGGG	5114
QY	5000	ACCGGGGAGCAAAATGTGTGTCTGAATCTTACCTTGAACCCGACTCTGTGCTACTAAAGGCA	5059
Dp	5115	ACCGTGGAGAGAAATGTGTCTCGAGACTTACCTTGAACCGGCTCTGTGCCACAAAGGCA	5174
QY	5060	CAGTGCAGAAATTTGTGATGATCACTTTTGAACCATTTTGAACAGGCAACAGCTGACT	5119
Dp	5175	CAGTGCAGAAATTTGTGATGATCACTTTTGAACAGTGTGAGACAGCCACCGGGGCT	5234
QY	5120	CTGCCCTGCCCCGTGECATCAAGTACATGTTTGACTTCTGTGATGAGCAGGCTGATAAC	5179
Dp	5235	CGGCTCTGCCCCGTGECATCAAGTACATGTTTGACTTCTGTGATGAGCAGGCTGACAGC	5294
QY	5180	ATGGCATTCATGACCCGACGCTCGGCACTACTGTGGAAGAGCAATGTGCTGCCCCGTGAGGT	5239
Dp	5295	GCCAGATCAGGAGACCCCGATGTGTGCGCACACTGTGGAAGAGCACTGTGCTGCCCCGTGAGGT	5354
QY	5240	TTTGGGTCAACATGATCAAGAACCCGAGTTGTGTGATCATTCATTAAGAACGATCA	5299

Db	5355	TCGAGGTGAATGTGATCAAGAACCCCGAAGTTCTGTTTGCACATCCAAGAAGCAGATCA	5414
Qy	5300	CAGACGCCCTGCTCTCTGTGTGGTGCTCAGACCTTCATGAGACTCTTGTCTCCAGCTCAGAGC	5359
Db	5415	CGAATGCTGCTGCTCGGTGGTAGCCAGACCTTCATGAGACTCTCTGCTCTACATCCGAGC	5474
Qy	5360	ACCGGCTGGGCAAGAGATCGCCCTCCCAACAAGCTGCTGTATTGCAAGAGATATCCCAAGT	5419
Db	5475	ACCGCTGGGGAAAGACTCGGCCCTCAACAACTGCTCTTAGCGCAMAGAACTCCCAACT	5534
Qy	5420	ACAAGAATTGGGTGAGAGGTAATTCTCACACATAGGAAAGATGCCAGCATCAGCGACC	5479
Db	5535	ACAAAGCTGGGTGAGAGGTAATTATGAGACANTGCAAAAGATGGCATTCATCAGCGACC	5594
Qy	5480	AAGACATGAACGCACTACCTGGCTGAGCAGATGCCGATGCAACATGATGAGTTCAACACCA	5539
Db	5595	AGGACATAGGAAGCCTTACCTGGTAGGAGATGCCGCCCTCCACGCCAGGACTTCAGCTCC	5654
Qy	5540	TGATGTGCACTCTCAGAGATCTTCTCCATGTGGGCAATPACAGGAGAGAGATCCTTGGAC	5599
Db	5655	TGATGTGGCTTCAAGAGACTGATTTCTATGTCAACAAAGTACCGCAGAGATTTCTCAGG	5714
Qy	5600	CTCTGGAACCAAGATGACCAAGTGTGGGAAGCAAACTGSCCTTACAACTGAAACAATCA	5659
Db	5715	CTCTGGACCGGAGATGCTTGTGTGGAAGATAAGTTGCGGCGAGAAACTGAAACAGATCA	5774
Qy	5660	TAACCTCATGAGCTTAGACAGCTGA	5665
Db	5775	TCAGCTGTGTTCCAGCGACAGCTTAA	5800

RESULT 11
LOCUS AX866475 5895 bp DNA linear PAT 29-MAR-2003
DEFINITION Sequence 31 from Patent WO2057450.
ACCESSION AX866475
VERSION AX866475.1 GI:29372168
KEYWORDS
SOURCE ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Edinger,S., Macdougall,J.R., Miller,I., Elleman,K., Stone,D.J.,
Gerlach,V., Grosse,W.M., Alsbroek,J.P., Lepley,D.M., Rieger,D.,
Burgess,C.E., Casman,S.J., Spytek,K.A., Boldog,F.L., Li,L.,
Padigerau,M., Mishra,V., Patturajan,M., Shenoy,S., Raestelli,L.,
Tchiernev,V.T., Vernet,C.A., Zernusen,B.D., Malynkar,U.M., Guo,Y.,
Miller,C.B. and Gangollis,E.A.
TITLE Proteins and nucleic acids encoding same
JOURNAL Patent: WO 02057450-A 31 25-JUL-2002;
CURAGEN Corporation (US)
FEATURES
source location/Qualifiers
1..5895 /organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 44.5%; Score 2533.8; DB 6; Length 5895;
Best Local Similarity 68.0%; Pred. No. 0;
Matches 3812; Conservative 0; Mismatches 1712; Indels 81; Gaps 17;

Qy	141	CGCGGAGGCTTTCATCACTCGGTGTGTGATGAGAGACAGACACATTTACTTGGGGC	200
Db	217	CGACTGGGGCTCCACCACTAGTGTCATGACACAAGCGCAGGTGTATGTGGCGC	276
Qy	201	CGTCAATCGATTATCAAGCTCTCCACGACCTGAAGTCTTGGTGAACGATGAGACAG	260
Db	277	AGTAAACCGCATCTATAGCTGTGGGAAACCTGACACTGCTGGGGCCCAACGTCAAGG	336
Qy	261	GCGGACGAGGACAAACCCCAAGTGTATACCAACCCCGCATGTCCAGACTGTGAATGAGCC	320

337 CCTGTGAGGAGCAACGAGAGTGTACCCGCGCCGAGCGTGTGACGTCGCCCCACGG 336
QY 321 CCTGACCAACCAACCAATGTCAACAAGATGCTCTCATAGACTACAGAGGAGGAGGCT 380
Db 397 CCTGGGAGTACTGACAAAGCTCAACAGAGCTGCTGCTGTGAGACTATCCGCTAACCGCCT 456
QY 381 GATTGCTGTGGAGCTGTACCAAGGATGTGCAAGCTGTGAGGCTGAGAGACTCTT 440
Db 457 GCTGCTGTGTGACGCGCTCCCAAGGAGTGTGCAAGTCTGTGCTGTGAGCAATCTCTT 516
QY 441 CAAGTGGGGAGGCTTATCATTAAGAGGACATATCTGTCAAGTGTCAAGAGAGCG 500
Db 517 CAACTGGGTGAGCCACACACCGTMAAGAGCACTAAGTGTCCAGCGCTGACAGAGGAGG 576
QY 501 CTCACTCTTTGAGAGTGTGTCTCTTACA-----GCAACTGATGACAACTGTTGAT 554
Db 577 CAGCATGGCGGGCGTGTCTATTGCGGGCCACCGGGCCAGGGCCAGAGCTCTTGTGT 636
QY 555 TGCCACGGGAGATGAGATGGGAGACCGGATTTTCCACCATCTCCAGCGGAAACTGAC 614
Db 637 GGGACACCCATCGATGGCAAGTCCGAGTACTTCCCACTGTCCAGCGCTGAGCTCAT 656
QY 615 CAGAAGCTCTGAGGCGGATGGCATGTTGCGAGTACGTCTTCCATGATGATGCTGTGCGCTC 674
Db 697 GGGCAACGAGGAGATGCGGACATGTTGCGCTGTGTGACAGAGATGATGTTGTGTATC 756
QY 675 GATGATTAAGATCCCTTGGGACACCTTACCATGATCCCTGACCTTGAATATCTACTATGT 724
Db 757 ACAGTCAAGATCCCTTGGACAGCTGTCCAGATTCGCGCTTGTGACATCTAATGT 816
QY 735 CTATGTTTATGACAGTGGCACTTGTGTACTTTTGAACCTCCACCTGATGATGATGTGTC 794
Db 817 GTACAGCTTCGCGACGAGAGTTTGTGTACTACCTCACGCTGACGTAAACACACAGCT 876
QY 785 TCCACAGGCTTCACACACAGAGAGAGGTGTATACATCCAGCTGTGAGGCTTTGCCA 854
Db 877 GACCTCGCTGATGCGCGCGGAGAGACTTCTTACGTCMAAGATCGTGGCGCTGTGT 936
QY 855 GGAGAGACAGAGCTTCAATCTCTATGTAGAGTGGCCATTGGCTGTGAGCGCAGTGGAGT 914
Db 937 GGAAGACCCCAATTTCTACTGTGATGATGATGATGATGATGATGATGATGATGATGAT 966
QY 915 GAGATGACGCTGTGCGAGGCTGTCTTACTGTCCAAAGCGGGGCGGCTGTGCGAGAC 974
Db 997 GAGATGACGCTGTGCGAGGATGCTGTACTGTGAGCGGGCGGCGGCTGTGCGAGAC 1056
QY 975 CTTGAGTGCATCCAGATGATGACTGCTCTTCAACGCTTCTTCAAGGCGCAGAGCG 1034
Db 1057 GTTGGGCTGTGAGAGAGAGAGAGAGTGTGTACTGTGTGTGTGTGTGTGTGTGTGTGT 1116
QY 1035 GAAAAATGAATCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1094
Db 1117 CCGCGTGAAGCCACMAAGAGATCAGACTGTGCTGTGTCAAGGCTCAGAGGCAATCAAGA 1176
QY 1095 CCGCATTAAGAGCGGCTGAGTCTTGTTCACGGGGCGAGGCGACGCTGACGCTGCGCTG 1154
Db 1177 GAAAGTTAAGAGCGGCTGAGTCTTGTTCACGGGGCGAGGCGACGCTGCGCTGCGCTG 1236
QY 1155 GCTCAAGTGAAGGATCCCTGAGAGAGTGTGCTTAAACATTAAGATGATGATGATGAT 1214
Db 1237 GCTGTCTCAACAGAGAGTGTGAGTGTGATCAATGCTGCTGTGAGATGAGATGAGATGAG 1296
QY 1215 TGGCGTGGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1274
Db 1297 CCGGAGAGACTTCAACAGAGCGCTGTGGGGGAGAGTCAACATTAAGAGGAGAGCGCGCTGT 1356
QY 1275 CACGAGAGACAGGAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1334
Db 1357 GGTGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1416
QY 1335 GGCCTTTGTGGGCAACAAAGTGGCAGGTGAAAGATCCGGGTGATGATGATGATGATGATGAT 1390

Db 1417 GGTATTCGCGGACAGGAGATGAGCGCGCATCCGAGATCTGTGTGAGACTCTTCAAAACC 1476
QY 1391 -----GGGCAACGCGCTTCAGATGAGACGATGAGGATGAGATGAGATGAGATGAGAT 1445
Db 1477 CCGTGGCGGCGCTGCGCTGTGCTACAGAGAGGCTGTGCGCGGAGGAGGAGCGCGCATCT 1536
QY 1446 CCGGAGATGAGCGCTTCTCCAGAGACACAGAGCACTTACATCATGATGAGAGAGAGCT 1505
Db 1537 GCGAGACTCTGTCTTCAAGCT 1596
QY 1506 CACAGAGTCCCTGTGAGT 1565
Db 1597 GACGCGGAGCTGTGAGAGAGT 1656
QY 1566 AGGCAACCGCT 1625
Db 1657 ACAGGAGCT 1716
QY 1626 TGAGCGGTCCAGAGAGCT 1685
Db 1717 TGAGGAGAGAGAGAGCT 1776
QY 1686 GGTCCATCCCAATATCTCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1742
Db 1777 TGTGAGCT 1836
QY 1743 GTACATGTCCCGGAGCTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1802
Db 1837 CTGAAAGT 1896
QY 1803 GATGAGGCTGT 1862
Db 1897 TGAAGCGGT 1956
QY 1863 CCGGATCA-----TCAAGAGATGGGAGACCAAGATGTGTGTGTGTGTGTGTGTGTGTGT 1916
Db 1957 GCGCATCGCGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2016
QY 1917 AAGAGAGCGGAGATGCTTGTGAGACAGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1976
Db 2017 CAGGAGACAGGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2076
QY 1977 CAATTG---TGCCTGT 2033
Db 2077 CAGTGTGAGT 2136
QY 2034 GATGTGTGACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2093
Db 2137 CCAAGTGTGACACACAGAGT 2196
QY 2094 CAGAGCTGCT 2153
Db 2154 GCTTACAGCTGAAAGCGCAAGACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2213
QY 2197 TGAAGCTGCCCAAGATCTGT 2256
Db 2257 ACCCATCACTTGGCGCGCACGAGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2316
QY 2214 ATGCAATCTCAATTCAG 2273
Db 2317 GTGCTCTTTCACATTCGAG 2376
QY 2274 CAGGATGAGTGCACAGACCTCTTATTCATGAGAGAGATGAGATCAACACCTGTGC 2333
Db 2377 CAGCTGTGAGTGCACAGATTCCTCTGTATCTTACAGAGAGAGAGAGATGATGAGAGAGAG 2436
QY 2334 CGTGAGTGTGAGAGT 2393
Db 2437 AGTGAACCTGTGAGT 2496
QY 2394 AGTTCACTTCACTCACTGT 2453
Db 2497 GCGGACCTCTCAAGT 2556

QY	2454	CCGAGACTTGGCATGTGGCTGTGTGCGAAGGCCCAAGGCCAATGTCACCTGTGGCCCAAGCACTG	2513
Db	2557	CCCGCGCTTCAGATGGGATGATGTGTGTGTGGAGCGCCGCTGTCTCCCTGGCAACCACTGT	2616
QY	2514	CCCTGCCCCA---GGAGAACCCAGTGGCTGGAGCTGTCTGTGTGTCCCAAAAGCAAGTGCACAAA	2570
Db	2617	CGTGTCCGACACACTGTGATCGTGTATGCAAGCGGTGTCAAGCCAGACAGATCGCTGCACCGA	2676
QY	2571	CCCCCGCATGCACAGATATATCCGGGTACAGAGCCCCCGGGGAAGAGGGGCACCAAGTCTAC	2630
Db	2677	CCCCAAGATCTCAAGCTGTCTCCCGGAAGGGGCCCGAGGCGAGGGCCGACCGGGCTCAC	2736
QY	2631	TATCCGAGGGGAGAACTTGGGCTGTGAAATTTTCGCGACATCGCCCTCCATGTCAAAGTTGC	2690
Db	2737	TATCAACAGCGAGAACTTGGGCTGTGCATTCGAAAGACTGTGCTGTGGCGTGGCGGTGG	2796
QY	2691	TGGCGTGAATGTCAGGCCCTTTAGTGGATGTGTACATCCCTTCGACAAACAGATCGTGTGA	2750
Db	2797	CAAGGTGTGTGCAAGCCCTGTGGAAGCGAGTACATGTCGCGAGCAAGATGTCTGTGA	2856
QY	2751	GATGGGGAGGCCAAGCCCC--AGCAGACATGACGGCTTCGTGAGATCTGCGTGTGT	2807
Db	2857	GATCGGGAGCCGACGCTCGTGGCGGCCCATGACGCTGTGTGAGGTGTGTGTGGGGA	2916
QY	2808	GTGTTCGGCTGAATTCATGTGCCCCGTCTCTACAGTCTATATCTTCAAGACATGACTCT	2867
Db	2917	CTGCTCAACACACTACCGGCCCCGTGTCAACCAAGGCTTCACTTGTGTGAACCAACCTT	2976
QY	2868	CTCAGATCTGAAGCCCAACCGGGGCCCATGTTCGAGAGGGAACCAAGTACCATCAAG	2927
Db	2977	CTACCGTGTAGCCCCCTCCCGTGGCCCTGTGTCAAGGGGCACTGTGATGTGCATCGAAGG	3036
QY	2928	CACCAACTGTATGCCCCGAAGCAAGTGTGTGTATGTTTGGAAAGCAAGCCCTGTCTT	2987
Db	3037	AAGCCACTGAACGAGGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3096
QY	2988	CCACAGCGGATCTCCATCTTACATTTGTGTGAAACACCAATCTTCAGATGAGTGTCTAGA	3047
Db	3097	CTCCGTGTCAAGAGAACTCCCGTGAATCCGCGTGAATCCGCGTGTGAACCCCCCGGAGAGGCC	3156
QY	3048	GATGAAGTGTTCGGTGCAGTGTGAGAGAGGCC-----AAGATCCACAGACCTGTGT	3098
Db	3157	TGGCAGCGCTCCCATCATCATCATCATCATCAACCGCGCCAGCTCACCAACCTTGAGGTAA	3216
QY	3099	CTTTCAGTATGTGAAGAAGCCCAACCATCGTGTGTGAGTTGAGCCAGATGAGACATTTGTGAG	3158
Db	3217	GTACAACTACACCGAGAGAGCCCAACCATCTGTGAGATGTACCCCGAGTGAACATCAAG	3276
QY	3159	TGGAAACACACCCATCGCGCGTATGGGGGACCCACCTGTGACCTCATACGAAGCCCGCAT	3218
Db	3277	CGTGTGGAGCCCTCTGACGATTCACAGGCAACCACTGGCCACTGTCTCGTGAACCCGAAT	3336
QY	3219	CCGTGCCAAGATGAGGAGGAAGACATCAATATCTGTGAGTTCTGAACGCTCATGA	3278
Db	3337	CCGGGCCAAGTATGAGGCAATGAG---AGGAGAACTGCTGTGTGTGAATGACACAC	3393
QY	3279	GATGACCTGTCAAGGGCCCCCGCTGTGTGGTCTGTGACCAACAGTACAGACTGACCGA	3338
Db	3394	CATGTATATGCCGCGCCCGGTCTGTGTGGCCAACTCTGTGGCAAGCCCAACAGACTGTGGGGA	3453
QY	3339	GAGGCCGAGAGATTGTGCTTCATCTGTGAACAACGTTCACAGTCCCTGCTCATCTCAACA	3398
Db	3454	GCGGCGGATGAGCTGGGCTTGTGTATGGAACAACGTGTGCTCCCTGTGTGTCAATCTC	3513
QY	3399	GACCAACTTCACCTATCCCAACCGGAGTTTAAAGCGTTTGTTCCTCAGAAATCT	3458
Db	3514	CACCTCTCTTCTTACCTGATCCCGTATCTGAAGCACTGACGCCCACTGGCGTCTCT	3573
QY	3459	GGAGCTCAAGCTTGCGACGCCCATCATCTCTAAAGGCAAGAACTGTATCCCGCTGTGGC	3518
Db	3574	GGAGCTGAAGCCCAAGCTCCCACTCATCTCTCAAGGGCGGAACTCTTGGCCACCT---GC	3630

OY	3519	TGGGGGGCAAGTGAAGCTGAACATCACTGTGCTGTGTTGGGGAGAAAGCCGTGCACCTGAC	3578
Db	3531	ACCCGGCACTCCCGACTCAACTCACTACCGGTGCTCATGGCTCCACAACCTGTACCTCTAC	3590
OY	3579	CGTGTCAATGATTCAGTGTCTCTGCGAGTCCCCCACTCATCTGGCAGGCAACAATGAT	3638
Db	3591	CGTGTCCGAGAACGACATCTGTGTGCGAGGGGCCCACTCATCTGGGAGCAACAAGTAC	3750
OY	3639	GGCCCTGTGCGATGGAGTACTCCCGGGATGTGTACATTGACCCCGGACAGCCC	3698
Db	3751	GTGTGCTGTGCAAGTGTGCTTGCAGTTCTGTCCAGGGAACCTGCAGGTGTAACTCCGACACCT	3810
OY	3699	GCTCAGCTGTGCCCCCAATGCTCAGCATGCGATGTGTGCGGCGCTCTCATATTTCAT	3758
Db	3811	GCTGACCTGTGCTGCTCCATTGTGGCCATTGGCGAGGGGGGGTCTCTGTGCTGTGTAT	3870
OY	3759	CGTGGCGGTGTCACTTCTCATTAACCGCAAGTCCCGGAGAAATGACTCAAGCTGAAGCG	3818
Db	3871	CGTGTGTGTGTCTATGCTCTCAAGCTTCTGTCCAGGGAACCTGCAGGTGTAACTCCGACACCT	3930
OY	3819	GCTGCAGATGCATGAGCAACTCTGAGTCCCTGTGTGTGCCCCCTGAGTGTCAAGAAAGCTT	3878
Db	3931	GCTGCACTCCAGATGGACAACCTGGAGTCCCGGTGTGCCCCCTGAAATGCAGGAAGGCTT	3990
OY	3879	TGCCGACTGCAAGCGACATCCATGAGCTGACACGATGACTCTGATGAGACCGGGATTTCC	3938
Db	3991	TGCAGACTGCAGACAACATCCACAGCTGTACCAATGACTGTGACGGTGTCCGCACTCC	4050
OY	3939	GTTCTCTGACTATTAAGACTTAACCATGTGCGGTGTCTTCCAGAAATTGAAGCCACCC	3998
Db	4051	CTTCTCTTACTACCGGACATATGCAACGCGGGTGTCTTTCTGTGGATTCAGAGCAACCC	4110
OY	3999	TGTCTCCGGGACCTTAGGTCCGGGCTACCGGCAAGACGTGTGAGAAAGGCTCGAA	4058
Db	4111	TGTCTCAAGAGATGAGAGTAAAGGCA-----TGTGAGAAAGTGTCTGAC	4158
OY	4059	GCTCTTGCCCACTTCAACAACAAGGTGTCTGTCTCTTCACTCCGACGCTTGA	4118
Db	4159	ACTGTTCGGGAGGTGTGTACCAAGAAAGCACTTCTGTCTACCTTCACTCCGACGCTGGA	4218
OY	4119	GTCCAGCGTAGTCTTCTCCATGCGGACCGTGTGCACGTGGCTCACTCATCATGACGT	4178
Db	4219	GGCAACGGGAGCTTCTCCATGTGGCGACCGGGAAATGTGGCTCTGCTCATCATGACGGC	4278
OY	4179	GCTGCAGACGACTGTGAGTACGCTCATGTATGTGTGAACAGTGTCTGCGGCACTCAT	4238
Db	4279	CTTCCAGGGCGAGATGGAATACGCAAGGGGTGTAAAGACTGTCTTCCGACCTCAT	4338
OY	4239	TGACAAAGACTGAGAGCAAGAACCAACCTAAGCTG---CTGCTCAGAGAGACTGATGTC	4295
Db	4339	CGAAGAGACTGTGAGGCAAGAACCAACCAACTGTACTGTGGCGGGCCAACTGAGTC	4398
OY	4296	AGTGGCTGTGAAGATGTGTACCAATTGTATTCTTCTCTCTCAACAAGTCCCTCAAGGA	4355
Db	4399	GGTGGCAGAGAAAGTGTACTAATCATGTGTCACTTCTCTTGTATAGTTCCTCAAGGA	4458
OY	4356	GTTGTGTGTGGAGGCCCTCTTCTCTCTGTCTGTGTCTGTGCATCAAGCAGCATGTAGAAAGG	4415
Db	4459	GTGGCTGTGGGAGCGGCTGTTCATGTGTGTACTGTGGCATCAAGCAGCATGTAGAAAGGG	4518
OY	4416	CCCCATTGACGCTATCACGGGCGAGGGCCGCTATTCTTGTAGCGAGCAACAAGCTCATCGG	4475
Db	4519	CCCCATTGACGCTATCACGGGTGTGGACGCTACTCTCTGTGTAGGCAAGAGCTCATCGG	4578
OY	4476	CCAGCAGATTGACTACAAACCTCTGTCTGTAGCTGTGTCAAGCCAGACAATGCGCAACG	4535
Db	4579	GCAGCAGATTGACTACAAACCACTGACCCCTGAATGTGTGTGAACCTGTGAATGTGAATGC	4638
OY	4536	CCCCGAGGTCCCAATTAAGATCCCTCAACTGTGTGACACATCACTCAGGTCAAGGAGAAAT	4599
Db	4639	ACCTGAAGTGTGGGTGAAGGGGCTGTGACTGTGTGAACGGTCAACCGGCTCAAGGAGAAAGCT	4699
OY	4596	TCTGATGTCCATCTTCAAGATGTGCTTGTCTCCACCGGCCCAAGCTGCAGATTGGA	4655

Dp	4699	GCTGACGCTGCTTACAGAGCGTGCCTTACTCCAGGGCCAGAGCCGCGAGCTGGA	4738
Qy	4656	TCTGGAGTGGCGACAGAGAGTGGGCGAAGGATGATCTTGACAGATGAAGACATACAC	4715
Dp	4759	CTGGAGTGGCGCGACGGGCGCGATGCGCCGCGATCATCTCTGCAGAGACGAGGACGTACAC	4818
Qy	4716	CAGATTTGAGAAATGATTTGGAGCGAGTGAACACACTGAGCCCACTACCAAGTGCAGATGG	4775
Dp	4819	CAGAGTTGACAGATTTGGAGAGAGGCTGAACACCTGGCTCATTCAGAGTACAGACGG	4878
Qy	4776	TTCCGTGTGGCACTTAGTGTCCAGACAGTGAACGCTTATACGAGTGAACAATCTCAC	4835
Dp	4879	GTCTCTGGGTGGCACTGTGTCCCAACACACTCCGCTTACACATCTCCAACTCTCTCAC	4938
Qy	4836	CGTCTCCAGGACCTTCAGCAATGTAATATGAAACATGATCCGGTACACCGGGCGAGCCCGA	4895
Dp	4939	CTTACACGAG---TCCCTCAGAGATGAGAGAGATCTCTGGACGCGCCAGACCCCGA	4995
Qy	4896	CAGCCTCCGCTACGCGACACTTATGATCATCTCTGACCTTGAAGTGGAGTCAAGTGTG	4955
Dp	4996	CAGCCTCCGCTGCGGACGCCCATATACGCCGACCTGGAAGAGGCGGACCAACTGTGG	5055
Qy	4956	GCACTAGTGAAGAAACACAGACGAGACCAAGAGAGAGGGAGCCGGGGGAGCAAGT	5015
Dp	5056	GCACTGTGTGAAGAAACACAGACACTGTGACCAACAGCTGAGAGGCTGACCGGAGCAAGT	5115
Qy	5016	GGTGTCTGAATATCTACCTGACCTCCGACTCTTGGCCACTA---AGGGCACTGCGAGAGTT	5072
Dp	5116	GGTCTGTGAGATCTACTTGAACAGGCTATCTGGCCACCAAGCAGGCGACACTGCAAGATT	5175
Qy	5073	TGTGATGACCTCTTTGAGACCATTTGAGACGCGACACCTGTGCTGCGCTCCCTCT	5132
Dp	5176	TGTGAGACGACTGTTTGAAGCACTTTAGGACGGCACACCGGAGGCTCAGCCTTGTGCT	5235
Qy	5133	GGGCTATCAGATACATGTTTGAATCTTCTGAGTAGACAGGCTGATTAACATGGCAATTGAG	5192
Dp	5236	GGGCACTCAAGTACATGTTTGAATCTTCTGAGTAGACAGGCGGACCAACACAGATCCAGAC	5295
Qy	5193	CCCCGACGTCGCGCACTACCTGGAGAGGCAATT---GGCTGCCCCCTGAGGTTTGGGATCA	5249
Dp	5296	TGCTGACGTGCGCGCACACTGGAGAGCACTGACGCTGCCCCCTGCGCTTGCGGTGAG	5355
Qy	5250	CATGATCAGAAACCCGCGAGTTTGTGTTGATCATCAATAAGAGAGATACAGACGCTG	5309
Dp	5356	CGTATCAAGAAACCCAGCTTGTGTGATTCATCAAGAAAGATCATCGAGGCGCTG	5415
Qy	5310	CCCTCTGTGTGGCTCAGACCTTCATGACCTTGTGCTCCAGCTGAGACACCGGCTGGG	5369
Dp	5416	CTGTGTGGGTGTGGCCCGACCTTCATGACTCTGTGCTCCACTGAGACAACCTGGG	5475
Qy	5370	CAGAGCTGCGCCCTCCACAGAGCTGTGATAGCCCAAGAGATATCCGAGCTAAGAAATG	5429
Dp	5476	CAGAGCTACCTTCACAGAGCTGTGATAGCCCAAGAGATATCCCACTTACAGAGCTG	5535
Qy	5430	GGTGA---GAGGATTTACTGACATAGGAGAGATCCAGCCATCAGGACCAAGACAT	5486
Dp	5536	GGTGGAGAGAGATCTACTATGACAGATGACCAAGATCCAGCATCAGCGACCAAGACAT	5595
Qy	5487	GAAAGCATCTCGGCGTGAAGAGTCCCGGATGACAGTAAGATGAGATTCACACATAGAGCG	5546
Dp	5596	GAGTGGATATCTGGCTGAGAGCTCCCGCTTACCTTACGCACTTCAACAGATGAGCGC	5655
Qy	5547	ACTCTCAGAGATCTTCTCTATGTGGCAATATACGGA-----GAGATCTTTGAGCC	5600
Dp	5656	CTTGCACGAGATCTACTCTCATCAACCAAGTACAGAGTGAAGTGCAGATCTTGGAGC	5715
Qy	5601	TTGAGACCAAGTGAACAGTGTGGAGACAGAACTGGCTTACAATCTTGAACAGATGAT	5660
Dp	5716	CTTGGAGAGAGATGAGAGGCGCGCGGCGAGCGGCTCGGAGCAACTGATGACAGTGTGT	5775
Qy	5661	AACCTCATGAGCTTAGACAGCTGA	5685

```

Db          5776  GGACACGATGGCCCTGACGACCTTA 5800

RESULT 12
LOCUS      HSN0VPROT
DEFINITION Homo sapiens mRNA for partial NOV/plexin-A1 protein.
ACCESSION X878932
VERSION    X878932.2 GI:6010216
KEYWORDS   NOV/plexin-A1 gene; transmembrane receptor.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS     Maestri, E., Tamagnone, L., Longati, P., Cremona, O., Gullisano, M.,
            Bione, S., Tamadini, P., Neel, B.G., Toniolo, D. and Comoglio, P.M.,
            A family of transmembrane proteins with homology to the
            MET-heptocyte growth factor receptor
            Proc. Natl. Acad. Sci. U.S.A. 93 (2), 674-678 (1996)
JOURNAL     96149362
MEDLINE     8570614
PUBMED      8570614
REFERENCE 2
AUTHORS     Tamagnone, L., Artigiani, S., Chen, H., He, Z., Ming, G.L., Song, H.L.,
            Chedotal, A., Winberg, M.L., Goodman, C.S., Poo, M.M.,
            Tessier-Lavigne, M. and Comoglio, P.M.,
            Plexins are a large family of receptors for transmembrane,
            secreted, and GPI-anchored semaphorins in vertebrates
            Cell 99 (1), 71-80 (1999)
JOURNAL     99449305
MEDLINE     10520995
PUBMED      10520995
REFERENCE 3
AUTHORS     Toniolo, D.
            Direct Submission
            Submitted (08-JUN-1995) D. Toniolo, Istituto di Genetica
            Biochimica ed Evoluzionistica, CNR, Via Abbiategrasso 207, 27100
            Pavia, ITALY
JOURNAL     Biochimica ed Evoluzionistica, CNR, Via Abbiategrasso 207, 27100
            Pavia, ITALY
REMARK      4 (bases 1 to 5263)
REFERENCE   Revised by [4]
AUTHORS     Tamagnone, L.
            Direct Submission
            Submitted (27-AUG-1998) L. Tamagnone, University of Torino, INCC,
            SP 142, 10060 Candolo, Torino, ITALY
JOURNAL     On Oct 4, 1999 this sequence version replaced gi:1247575.

COMMENT
FEATURES
Source
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /chromosome="3"
    /map="q21-qter"
    /tissue_type="skeletal muscle"
    /clone_lib="lambda GT10"
    /gene="NOV/plexin-A1"
    .1..5263
    <1..>5263
    /gene="NOV/plexin-A1"
    /codon_start=2
    /product="NOV/plexin-A1 protein"
    /protein_id="CA57274.1"
    /db_xref="GI:6010217"
    /db_xref="GOA:Q9UW2"
    /db_xref="SPTREMBL:Q9UW2"
    /translation="GMAEAGLFRAGGSGQPEPFTTSGSDMGLTHLLVHBOGEVYVG
    AVNIVYKSGNTLIRAHVYGVPEDEKCYPPYQSCPHGAGNTDWNKLLILDVAA
    NRLACASAGQICQSLRLDXLFKLGEPHREKHTLSVQAGSMAGVLTAGPGQGO
    AKLVNPIIDKSEYFPTLSRLANERDAMGFVYQDSFSSQLKISDLSKRP
    AFDLYVYSPRSBOFYVYLLQDITQTPDAAGHPTSKIVRLCYVDPEYVYVF
    PIQEGQAGVEYRLVDQAVLSRREGALAHQGLAEDEDVLFVPAQGGKNRKPKEEA
    LCLTIAIAIKETIKERIKQSCYRGEGCLSPMLINKEIGLCINSPICDDPRGDFNPG
    LGGVTIIGTLPFNDKDGITAVAAIDYHGRVYVYAGRSGRIRIKLYLDSNPEGRA
    LAYSVAGSGSPILKDLVLSNPHQTLVAMTEKQVTRVPSVCQVYTSCELCASSGP
    HCGVCAVHSGICSSRDACERADQEPAAADLLQCVLTQVPPNVAIVMSQVEPLVQANN

```

VEDLAGVNCSEFDETESESVLEDRHCRSHSAREVAPITRQGDQKRVKLYKXKKE
 TEKKFASVDVFFVNCSSVHOSCLSCVNGSFPCMKRYRCHTNVADCFLEGRVNVSE
 DCPOLIPSTQIVPVGVKPIITLAAKNLPPQSGQGECEFLHFGSAPRTALRPN
 SLOOCSNYSYENGVDSPVNLVNMVNMVNDIPNIOAHLYKCALRESGLCL
 KADPREFCQMCVABRCSLRHCHCADPTASMMHARSGSRCPDPLIKLSPEPCOG
 GTRILITGNGNLARPEDVRLGTRGVKULCSPESEYIABOIVCEIGPSSVRHDL
 VAVCRDCSPHRLALSPKSFVITPTFRVSPSRGPGGWTGILBGLHAGSVAY
 SVGAPCSFWSNRSRIRCLTPPGQSPSAPILININAQOLNPKVCEKNTYEDPTILR
 IDPWSINGEGLTLVTGTNLATVREPRIRAKYGGIENEGCLVNDTPTGLKPSVA
 NPYRSPDELGFMDVNRSLVYNSTFLYEDPPLSPGLLEKPSF
 LILKGNLIPRAGNSRLVYTLQSTPCTLTSETOILCEAPNLPGQHTVAGAGL
 EESPTLOYSDSLTLPAIVGIGGGGGLILLVAVLILAKRSRDADEPTLKLQLO
 MNLEBRVALBECKEAPAEIQTIIHNDLBDAGIPFDIYRTYMRVYFPEIDHPPL
 KMEVGVANVKSJLTFQULIKKHFLLFIRTLKORFSMEKRNVAJLIMTALQGS
 MEYATGVLKQLISDLIERULESKNHPKILRATESVAERKLNMTFLLYFLKACG
 EPLFLMYCALKOOMEKGPIDA.TGEARVSLSEDKLIRQIDYKILTLNCPENENAP
 EYVKGJLDCDVTQAKERKLDAAYGVAVSPORPKAADMLEWROGMARILLOEDVY
 TKIDNMKRLNLTALHOVTDSSVALVPRKOTSANISSNSTKSLSEYSEMLRTAS
 PDSLRSPMTIPDLSEGTKMLHLYKNHLDORRGGKXVSLVYLTLLATKGLTL
 QKRFDDIPEITISTAHRSALPLAKNMFDFLBDQADHGQIHDAVDRITWMSNCLPLR
 FVNVIKNFQFVFDLHKOSITDACLSV"

ORIGIN

Query Match 42.6%; Score 2423.2; DB 9; Length 5263;
 Best Local Similarity 67.7%; Pred. No. 0;
 Matches 3525; Conservative 0; Mismatches 1639; Indels 42; Gaps 8;

QY 141 CCGCGAGGGTTTCATCACCTGGTGTGATAGAGAGACAGACATTTACTGGGGGC 200
 DB 76 CGACTGGGGCCTCACCCACCTATTGTGTGATAGAGAGAGCGAGGTGTATGTGGGGGC 135
 QY 201 CGCATATGGATTATCAAGCTCTCCAGCAGCAGGTCTTGAGAGCATAGAGACAG 260
 DB 136 AGTGAACCCGATTAAGTGTTCGGGGAACCTTGACATCTGTCGGGCCCACTACCGG 195
 QY 261 GCCGAGCAGAGCAACCCCAAGTGTATCCACCCCGCATGTCAGAACCTTGACATGAGCC 320
 DB 196 CCTGTGAGAGCAACGAGAAGTGTACCCGCCGCCANCTGACAGTCTTGCCCCCAAG 255
 QY 321 CCTGACACCAACCAAAATGTCAAGATGCTCTCTCATAGACTACAGAGAAACAGGCT 380
 DB 256 CTTGGGCAATACTGACAAAGCTGCTGCTGCTGACATGCGGCTAACCGGCT 315
 QY 381 GATTGCTGTGGAGCCTGTACCAAGCATCTGCAAGCTGTGAGGCTTGAGACCTCTT 440
 DB 316 GCTGGCTGTGGAGGCGCTCCAGGGCATCTGCGCATGTCGCGCTGGAACATCTCTT 375
 QY 441 CAAGCTGGGGAGCCTTATCATTAAGAGAGCATATCTGTCAAGGTGTCAAGAGAGG 500
 DB 376 CAAACTGGGTGAGCCACACACCGTAAAGAGCATACGTGTCAAGGTGTCAAGAGAGG 435
 QY 501 CTCAAGCTTTGAGATGATGCTCTCC-----TACAGCAACTGAGTGAACAAGCTTTGAT 554
 DB 436 CAGCATGGGGGGGTGTCTATTGCGGGCCACCGGGCCAGGGCCAGCAAGCTCTTGGT 495
 QY 555 TGCACAGGAGTGTGAGAACCCGAGTATTTTCCACCATCTCCAGCCGAAACTGAC 614
 DB 496 GGGCAACACCATGATGAGGCAAGTCCGAGTACTTCCCAACATGTCAGGCGGTGAGTCA 555
 QY 615 CAAGAATCTGAGCGGAGTGGAGTGTTCGGTAAAGCTTCATGATGATGATGATGATGATG 674
 DB 556 GGGCAAGAGAGAGATGCCACATGTTCCGCTTCGTGACAGATGATGATGATGATGATG 615
 QY 675 GATGATTAAGATCCCTTCGACACCTTACCATCATCTCTGACTTGTATATTAATGT 734
 DB 616 ACAGCTCAAGATCCCTTCGACACGCTGTCAAGTCCCGGCTTTGACATCTACTATGT 675
 QY 735 CTATGATTTAGAGTGGAGCAACTTGTACTATTTTGAACCTTCAACCTGAGATGATGTC 794
 DB 676 GTACAGCTTCGAGAGAGAGATTTGTCTACTACTCAAGCTGTGAGCTTGAACACAGT 735
 QY 795 TCCACCAAGGCTCCACCAACAGAGAGAGGTGTATACATCAAGCTGTGAGGCTTTGCA 854

DB 736 GACCTGCTGTATCCGCGCGAGACACTTCTTACGTCACAGATGTGTGCGGCTGTGTGT 795
 QY 855 GAGAGACACAGCCTTCAACTCTTATGTAGAGTGGCCATTTGCTGTGAGCGCAGTGGGT 914
 DB 796 GAGACGACCCCAATTTCACTGTACGTGTAGTTCCTCATTTGGCTGTGAGACCGGGTGT 855
 QY 915 GAGATACCGGCTGTGTAGGCTGCTTACCTGTCCAAAGCGGGGGCGGTGTGTGGCAGAC 974
 DB 856 GGAATACCGGCTGTGTAGGATGCTTACCTGAGCCGGCGCGGTGTGTGTGTGTGTGTGT 915
 QY 975 CTTTGAAGTCCATCCAGATATGACCTGTCTTACCGTCTTCTCCAAAGGGCCAGAGCG 1034
 DB 916 CTGTGGCTGTGTAGAGACAGAGACGTGCTTCAACGTGTGTGTGTGTGTGTGTGTGTGT 975
 QY 1035 GAAATTAATTCCTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1094
 DB 976 CCGGTGAGACCCCAAGAGAGTACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1035
 QY 1095 CCGCATTAAGAGAGGCTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1154
 DB 1036 GAGATTAAGAGAGGCTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1095
 QY 1155 GCTTCAAGTGAAGACATCCCTGTGACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1214
 DB 1096 GCTGTCAACAGAGAGT 1155
 QY 1215 TGGCTGTGACATGAATGCTCTCCCTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1274
 DB 1156 CCGGAGAGATTTCAACAGACCCCTGTGGGGGACAGTACCATGTAGAGAGAGAGCCCTGT 1215
 QY 1275 CACGAGAGACAGGACCCGATGACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1334
 DB 1216 GTGTGACAGAGATGATGCTGTGACCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1275
 QY 1335 GGCCTTTGTGGGACCAAAAGTGTGACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1390
 DB 1276 GTTATTTGCGGACCGGAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1335
 QY 1391 -----GGGGCAAGCCTCTCCAGTATGAGACGTTGAGAGTGTGTGTGTGTGTGTGTGT 1445
 DB 1336 CCGTGGCGGGCTGT 1395
 QY 1446 CCGGATATGAGCCTTCTTCAAGAGACAGAGCACTTATCATGTGTGTGTGTGTGTGTGTGT 1505
 DB 1396 GCGAGACCTGTCTTCAAGCCCAACAGTACCTTATGATGATGATGATGATGATGATGATG 1455
 QY 1506 CACGAGTCCCTGTGTGAGT 1565
 DB 1456 GACCGGGGT 1515
 QY 1566 AGGAGACCCCACTGT 1625
 DB 1516 ACGGAGACCCCACTGT 1575
 QY 1626 TGAAGGTCGAAGAGAGCCCGGAGGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1685
 DB 1576 TGAAGGTCGAAGAGAGCCCGGAGGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1635
 QY 1686 GTTCCATCCCAACATATCTCCGTCTGTCAATCAAGTGTGTGTGTGTGTGTGTGTGTGTGT 1745
 DB 1636 TGTGACAGCCCGCAATGT 1695
 QY 1746 CAATGTCCCGAGGT 1805
 DB 1696 GAACTGTCTGACCTTCAAGT 1755
 QY 1806 TGGGCTGT 1865
 DB 1756 GAGGCTCTGTGAGAGATGCGGAGATCACTGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1815
 QY 1866 GATCATACAGAGATGAGAGACCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1925
 DB 1816 CATCACCGGGGCGAGAGAGACAGCGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1875

QY 1926 CGGCATGACCTTCCGACGACCAAGCTTGTCTTCTTACATTTGAGAGCTTCCAAATTCGTG 1985
DB 1876 AGGGAAGAAGTTTGGCTGTGTGAGACTTGTCTTCTTACATGAGAGCTTCCACCAAGTCTTG 1935
QY 1986 CTTGTCTGCTGAGAGAGTCCATACCGCTGCATCTGTGTAAATACCGGAGATGTCTGAC 2045
DB 1936 CTTGTCTGCTGAGAGAGTCCATACCGCTGCATCTGTGTAAATACCGGAGATGTCTGAC 1995
QY 2046 CCATGACCCCAAGACCTTCTCTTCCAGGAAGCCGAGTGAAGCTGCCAGAGACTGCC 2105
DB 1996 ACACACTGTGAGTACTGTGCTTCTGTGAGAGGCGGTGTCAACGTGTCTGAGAGACTGCC 2055
QY 2106 CCACTGTGCTGAGTGAACAAGTCTGTGTGCTGAGAGTGAATCAAGCTTATCAAGCT 2165
DB 2056 ACAGATCTGCTCCCTCCACGACGATCTAGTCCAGTGGAGTGTAAACCCATCAAGCT 2115
QY 2166 GAAAGGCCAAGACCTTCCGACCCCAAGCTTGGGAGGCTGCTACGATCCATCTCA 2225
DB 2116 GAGCCGACGAACTGACCAACGACCAAGTCAAGCCAGCTGAGATATGAGTCCCTTCCA 2175
QY 2226 CATTGAGGACGAGACGACGAGTCCGCTGTGCTTCAACAGCTTCAAGCTTCAAGT 2285
DB 2176 CATCCGGGACAGCCCGGCTGTGTACCGCTGTGCTTCAACAGCTTCAAGCTTCAAGT 2235
QY 2286 CCAAGAACACTTCTTATTTCTTATGAAAGGATGAGATCAACAACGTCGCTGAGTGTAC 2345
DB 2236 CCAAGAACTCTGTACTCTCTACGAGGAGACGATGTACGCACTTGCAAGTGAACCTGT 2295
QY 2346 AGTGTGTGATGAGGCACTTCAACATGTACCAACCACTGAGATTAAGTTCACTCTA 2405
DB 2296 AGTGTGTGAGAACGCACTTGTCTTATGACCAACCAAGAACATCCAGGCGCACTCTA 2355
QY 2406 CAACTGTGAGACATGCTGTGAGAGTGTGGGCTGTGCTCAAGCTGACCCAGACTGCG 2465
DB 2356 CAAAGTCCCGGCTGCGCGAGAGCTGTGCTGTCTCAAGGCGCAACCGCTTGTGA 2415
QY 2466 ATGTGTGTGATGAGGCGGCGGCGGCAAGTGAACCTGTGCGGAGCACTGCCCTGCGCA 2522
DB 2416 GTGGGATGT 2475
QY 2522 GAGAGACCAAGT 2582
DB 2476 ACCTGATGT 2535
QY 2582 AGAGATTAATCCGAGTGAACGAGCCCGGAGAGGAGGAGCAACAAGTCACTATCCAGAGG 2642
DB 2536 CAAAGT 2595
QY 2642 GAACTGT 2702
DB 2596 GAACTGT 2655
QY 2702 CAGCCCTTATGAGT 2762
DB 2656 CAGCCCTGT 2715
QY 2762 CAAAGCC---AGCAGCATGACAGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2819
DB 2716 CAGCTGT 2775
QY 2820 ATTATGAGGCGGCTGT 2879
DB 2776 CTACGCGGCTGT 2835
QY 2880 GCCCAGCCGAGGAGCCATGTCCGAGAGGACCAAGTGCATTCACAGGACCAACTTGA 2939
DB 2836 CCGCTCCGAGGAGCTGT 2895
QY 2940 TGCCTGAGAGAAAGT 2999
DB 2896 CGCAGGAGT 2955

QY 3000 TTCACTCTTACATGT 3059
DB 2956 CTCCGTGAGATTCGGGTCTGT 3015
QY 3060 GGTGTGAGT 3116
DB 3016 CATCAACATCAACGCGGCGGAGCTTCAACCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3075
QY 3117 CCCCACATGT 3176
DB 3076 CCCCACATGT 3135
QY 3177 CGTATGAGGAGCCCACTGT 3236
DB 3136 GGTGTGAGGAGCCCACTGT 3195
QY 3237 GAAAGACATCAATATGT 3296
DB 3196 CATGTGAGGAGGAGAACGCTGT 3255
QY 3297 CGCCCTGT 3356
DB 3256 GT 3315
QY 3357 CTTATCTGT 3416
DB 3316 CTTGT 3375
QY 3417 TCCCAACCGGAGT 3476
DB 3376 CCTGT 3435
QY 3477 GCCCATCATCTTAAAGGAGAACCTGTATCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3536
DB 3436 CCCATCATCTTCAAGGCGGAGAACCTGT 3492
QY 3537 GAACTTACTGT 3596
DB 3493 CAACTTACTGT 3552
QY 3597 GCTGT 3656
DB 3553 GCTGT 3612
QY 3657 GAGTACTCCCGGAGATGT 3716
DB 3613 CGAGT 3672
QY 3717 CGTCAAGCATGT 3776
DB 3673 TGT 3732
QY 3777 CTATTAACGCAAGTCCCGGAGATGT 3836
DB 3733 CTACAGCGCAATGT 3792
QY 3837 CAACTGT 3896
DB 3793 CAACTGT 3852
QY 3897 CATTCATGAGT 3956
DB 3853 CATTCATGAGT 3912
QY 3957 TTACACATGT 4016
DB 3913 ATATGCAATGT 3972
QY 4017 GGT 4076
DB 3973 GGT 4020
QY 4077 CAAACAAAGT 4136

Db 4021 GACCAAGAGACCTTCTGCTGACCTTATCCGACGCTGAGGACACAGCCACCTTCTC 4080
 Qy 4137 CATGGGACCGTGGCAAGTGGGCTCACTGATCATGACCGTGGACAGAGAGCTGGA 4196
 Db 4081 CATGGGACCGGCGGAAATGGGCTTGGCTCATGACGGCCCTGGAGGGCGAGATGGA 4140
 Qy 4197 GTACGCACTGATGTGTGAAGCAGCTTGGCCGACCTCATGTAACAAGAACTTGAAG 4256
 Db 4141 ATACGCCACAGGCGTGTCAAGCAGCTTCTTCCGACCTCATGAGAAAGAACTTGAAG 4200
 Qy 4257 CAGAGAACACCTTAGCTGTCTGTCTAGAGACATGAGTCACTGTAGAGAGTCTGAC 4316
 Db 4201 CAGAGAACACCCCAAGCTGTCTGTCTAGAGACATGAGTCACTGTAGAGAGTCTGAC 4280
 Qy 4317 CAATTGTTTACTTCTTCTCTTCTTACAGTTCTTCAAGAGTGTCTGTGGAGCCCTCTT 4376
 Db 4261 TAACTGGTTACCTTCTTCTTGTATTAAGTTCTTCAAGAGTGTCTGTGGAGCCCTGTT 4320
 Qy 4377 CTCCCTGTTCTGTGTCTATCAAGCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 4436
 Db 4321 CATGTGTATCTGGCCATCAAGCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4380
 Qy 4437 CAGAGCCGCTACTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4496
 Db 4381 TGAGGCAAGCTTACTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4440
 Qy 4497 CCTGTCTTGAAGTGTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4556
 Db 4441 ACTGACCTTGAAGTGTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4500
 Qy 4557 CCTCACTGTGACACATCACTCAGGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4616
 Db 4501 GCTGACCTGTGACAGGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4560
 Qy 4617 TGTGCTTGTCTTCCACCGGCCCCAAAGCTGAGATATGAGATGAGAGAGAGAGAGAG 4676
 Db 4561 CGTGCCCTACTCCAGCGGCCCCAAAGCTGAGATATGAGATGAGAGAGAGAGAGAGAG 4620
 Qy 4677 TGGGCGAGAGATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4736
 Db 4621 CATGCGGCGATCACTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4680
 Qy 4737 GCGACTGACACACTGCGCCCACTACAGAGTGTCTGAGTGTGAGATGATGTC 4796
 Db 4681 GAGGCTGACACACTGCGCCCACTACAGAGTGTCTGAGTGTGAGATGATGTC 4740
 Qy 4797 CAGAGAGTGAAGAGCTTAAAGCAGATGACAACTTCCACGCTTCCAGAGAGAGAGAG 4856
 Db 4741 CAGAGAGAGTGTGCTTCAAGATGATGATGATGATGATGATGATGATGATGATGAT 4797
 Qy 4857 TAAATATGAAAACATGATCCGGTACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4916
 Db 4798 CAGATATGAGAGAGATGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4857
 Qy 4917 TATGATCACTCTGACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4976
 Db 4858 CATGATCAAGCGCCGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4917
 Qy 4977 GCAGGAGAGACAG 5036
 Db 4918 CCACTGAG 4977
 Qy 5037 CCGACTCTGAGCACTAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5096
 Db 4978 AGGCTACTAGGAG 5037
 Qy 5097 CTTCAG 5156
 Db 5038 CTTCAG 5097
 Qy 5157 CCTGTATGAG 5216

Db 5098 CCTGTATGAG 5157
 Qy 5217 GAGCAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5276
 Db 5158 GAGCAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5217
 Qy 5277 TGACATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5322
 Db 5218 CGACATTCACAG 5263
 RESULT 13
 AK122289
 LOCUS Mus musculus mRNA for mKIAA0463 protein. 6893 bp linear ROD 15-MAR-2003
 DEFINITION AK122289
 VERSION AK122289.1 GI:28972234
 KEYWORDS FLI CDNA.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCES
 AUTHORS 1
 TITLE Okazaki, N., Kikuno, R., Ohara, R., Inamoto, S., Aizawa, H., Yuasa, S., Nakajima, D., Nagase, T., Ohara, O. and Koga, H.
 JOURNAL Prediction of the coding sequences of mouse homologues of KIAA
 REFERENCE KIAA-homologous cDNAs identified by screening of terminal sequences
 ATTORNS of cDNA clones randomly sampled from size-fractionated libraries
 TITLE DNA Res. 10, 35-48 (2003)
 JOURNAL 2 (bases 1 to 6893)
 Okazaki, N., Kikuno, R., Nagase, T., Ohara, O. and Koga, H.
 DIRECT SUBMISSION
 Submitted (07-FEB-2003) Hirosaki Koga, Kazusa DNA Research
 COMMENT Institute, Laboratory for Genome Informatics, 2-6-7
 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
 (E-mail: mouse@kazusa.or.jp, Tel: 81-438-52-3919, Fax: 81-438-52-3918)
 The CREATE program supported by Japan science and technology
 corporation; cDNA full insert sequencing; Kazusa DNA Research
 Institute; cDNA library construction, clone selection and 5'-
 3'-end one pass sequencing.
 FEATURES
 source
 1..6893
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="mb907539"
 /isue_type="adult"
 /dev_stage="adult"
 /note="vector:modified pbc SK+"
 1..6893
 /gene="mKIAA0463"
 <1..4592
 /gene="mKIAA0463"
 /note="CDS is predicted by in silico analysis. Start codon
 is not identified."
 /codon_start=3
 /evidence="not_experimental"
 /product="mKIAA0463 protein"
 /protein_id="BAC65571.1"
 /db_xref="GI:28972234"
 /translation="OIKERLDSCHYGEENLEINWLLGKDVOCRTKAPVPIDNFCGADI
 NQPIGSGTPEVRLTYLTSRDLRTSVASVYVNGSVVPTGSGKLRKIRADGPFGG
 VQTEKSVFKDSPIRLDMARFINOYLTVMEKRVYVPTVPSCEQYTKRAGCSLSDG
 PFGWALNMSKRDCKORAEARFASISQCSLEVHPISISYSDSRLLSLYVN
 DAVNLSEGLACAGNLTVEGVSSQVCLISPGEDVAVLPDDWRELEQLKSKX
 TKRIFVSTEFKFNCSAHLCSVNSARCMKXRNLCITDPPTGCSFQERINVS
 DCPQLVTEIILIPVEVAPILTKANLEPQSGRGVCEVLSIQAVHVALRNS
 SSVQCNSSVQYDMDISNLAVDPVAVNAGNLIINPDQLKVALYCAQARSQGL
 KADHKREGCWGGERCTLHOCPSPTSMWMSHNVCSNPQITEILITVGGPPEG
 TVRTIGVNLGIDPFEIHHNOVAVCPPIGELYIARQICEMGHAVITGSGPVR
 LCTGCKPFPMTKSHQITFFVNPVSLSPISPIGSGGCTMTTITGHYIAGSSVAVYI
 GNQTCFEPERSMNEIVCSPSSNGLPVPSVSDRAVDSLOPEYITDPRVQXIE

Db 1863 GCAGTGAACCTGGGCAATCAGACCTGTGAATTCATATGAGGAGATGATGATTTGTA 1922
QY 3016 TG---CAACACACATCTTCAGATGAGGTGCTAGATGAAGGTGTCGTGACAGTGGAC 3072
Db 1923 TGTGTTTCAACCCCATCATCAATGAGACTAGAGCCAGTCCCTGCTCCGTAGAGTTCGAC 1982
QY 3073 AGGGGCAAGATCCACAGAGACCTGTCTTCATGATGAGAGAGCCACCATGTCGCG 3132
Db 1983 AGAGCCCGGTGATAGCAATGTCGATGTTGAGATATAGTACCCAGCGGTCCAACT 2042
QY 3133 ATTGACCAAGATGAGCATTTGTCAGTGAACAACACCATCGCCGATGAGGAGACCCAC 3192
Db 2043 ATTGACCAAGATGAGATATCACTAGTGGGACACACCCCTAACCATCAAGGCTTCAAC 2102
QY 3193 CTGACCTCATACAGAACCCCAAGATCCGTCCAGATGAGAGGAGAGACATCAAT 3252
Db 2103 TTGATGTCATTCAGAGACCCAGGATCCGAGTCAAAATTTAATGCAAGATCTGCAAT 2162
QY 3253 ATCTGTGAGTTCAGAACGCTACTGAGATGACCTGTGACGCGCCGCTCGCTCTGGGT 3312
Db 2163 GTATGCAAGTGTGAACAACAACACCTTCACTGTGAGACCCCTCTGACCAAGTAC 2222
QY 3313 CTGACCAACAGTCAGACTTGAACGAGAGGCCCAAGATTTGGCTTCATCTGGAACAC 3372
Db 2223 TACCGCCCAAGGTCTGGAACATGTTGAAGCGGCAAGATGATTTCTTTAACAAT 2282
QY 3373 GTCCAGTCCCTGCTGATCTCAACAGACCAACTTCACTTATCCCAACCCGCTTTT 3432
Db 2283 GTTCAATCTTACTATCTATGAACACCAAGTTTCACTTACCCCAACCCCAAGTTT 2342
QY 3433 GAGGCTTTGTGCTCCCTAGGAATCTGTGAGTCAAGCTGACGCCCATCATCTTAAAG 3492
Db 2343 GAACGTCTCAGGCCCACTGGAATCTTGGATCAGAAAGCCAGGCTCAACCATATCTTAAG 2402
QY 3493 GGCAGGAACCTGATCCCGCTGTGGCTGGGGCAACGTAAGCTGAACCTACATCTGTGTG 3552
Db 2403 GCGAAATATCTGTCTCTCTGCTCTGAGAG--GCGCAATCTCACTACACGTAATG 2459
QY 3553 GTTGGGAGAGAACCTGTGACCGTGAACGATGTCAGATGTCAGCTGCTGTGAGTCCCC 3612
Db 2460 ATTGAAGAGACACTTGAACAGTCACTGTGTGTAAGACACAGCTGCTTGAACCTCCC 2519
QY 3613 AACCTCATGAGGAGCAAAAGTATGAGCCGCTGTGCTGCTGAGATTAATCCCCGGGG 3672
Db 2520 AACCTCAAGGAGCAAAAGTATGAGTTCAGTGGGCGAGATGATGTCACCTGAC 2579
QY 3673 ATGTGTACATGCCCCGAGACAGCCGCTCAGCTGCCCCCATCTGTCAGATGCGAGTG 3732
Db 2580 TCCGTGAGGTATCTCCGACAGCTGTGTAACCTGCGACCATCATCAGCATGCGAGCT 2639
QY 3733 GCTGGCGGCTCTCTCATCTTTTCAATCGTGGCCGCTGCTCATTTGCTATTAACCGAATCC 3792
Db 2640 GGTGAAGGCTCTCTTATCATCTGTCATCTTGTCTCATCTGCTTAACCGCAAGCT 2659
QY 3793 GCGGAAGTGAACCTCAGCTGAAGCGGCTGAGATGAGTGAACAACCTGAGTCCGT 3852
Db 2700 AAGGAATATGACTCATCACTCAAGGGGCTCCAAATGCAATGGAACAACCTGAGTCCAGG 2759
QY 3853 GTGGCCCTGAGTGAAGAGACCTTTGCGAGCTGACAGAGCAATCATGAGTGAAC 3912
Db 2760 GTGGCACTGAGTGAAGAGACCTTTGCGAGCTTCAAGACAGATCAATGAGTGAAC 2819
QY 3913 AATGACCTGAGTGAAGCGGAGATCCGATCTGAGACTATAGAACTTAAACATGCGGGTG 3972
Db 2820 AATGACTTGAATGATCAGAAATCCCTTACCTGAGACTACGTAACCTAAGCACTGAGATC 2879
QY 3973 CTGTTCCAGGAATGAAGCAACCTGTCCTCCGGAACCTTGAAGTCCCGGCTTACCGG 4032
Db 2880 CTGTTCCAGGCAATGAAGCAACCTGTTCTGCGGGAACCTGAGATCAGGGAATGGA 2939
QY 4033 CAGGAGCTGTGAAGAAAGGCTTGAAGCTTTGCGCCAGCTCATCAACAAGGTGTC 4092

Db 2940 CAGCAGCAGTGAAGAAAGCCCTGAACCTTTCGCCAGCTTATCAACAAGGTGTC 2999
QY 4093 CTGCTGTCTCTTCAATCCGCAAGCTTGAAGTCCAGAGTACTCTTCATGCGGACGTGAC 4152
Db 3000 TTGCTGACCTTCTATCCGTAAGTGAACCTGGAACCTCAAGCAAGCTTCTCATGAGACCTGAG 3059
QY 4153 AACGTGCTCACTCATGACCGTGTGCAAGAGCAAGCTGAGTACGCACTGATGTG 4212
Db 3060 AACGTGCTCTCTCATGACAGGCTTCAAGGCTGAGGCTTGAATATGCACTGATGTG 3119
QY 4213 CTGAAGCAAGTCTGCGGACCTCATGTAAGAACTTGAAGAGAGAAACACCTTAAG 4272
Db 3120 CTGAAGCAAGTCTGCTGACCTCATGTAAGAACTTGAAGAGAGAAACACCTTAAG 3179
QY 4273 CTGCTGTGAGAGACTGAGTCAAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTG 4332
Db 3180 CTGCTTCTCCGAGAGACTGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTG 3239
QY 4333 CTCTCTTACAGTCTCTCAAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTG 4392
Db 3240 CTCTTACAGAGTCTCTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTG 3299
QY 4393 ATCAAGCAGAGATGAGAGAGGCGCCATGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCA 4452
Db 3300 ATCAAGCAGAGATGAGAGAGGCGCCATGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCA 3359
QY 4453 TTGAGGAGAGCAAGTCAATCCGCAAGATGTAATGTAATGTAATGTAATGTAATGTAATG 4512
Db 3360 CTGAGTGAAGAGAGTCAATCCGCAAGATGTAATGTAATGTAATGTAATGTAATGTAATG 3419
QY 4513 GTCAAGCCAG 4572
Db 3420 GTCAAGCCAG 3479
QY 4573 ATCACTCAGTCAAGAGAGAGATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 4632
Db 3480 ATCACTCAGTCAAGAGAGAGATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 3539
QY 4633 CGGCCCAAGCTGCAAGTGTGATCTGAGTGGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4692
Db 3540 CGGCCCAAGCTGCAAGTGTGATCTGAGTGGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3599
QY 4693 TTGCAAGTGAAGAGATCAACCAAGATGTAATGTAATGTAATGTAATGTAATGTAATG 4752
Db 3600 TTGCAAGTGAAGAGATCAACCAAGATGTAATGTAATGTAATGTAATGTAATGTAATG 3659
QY 4753 GCCCACTCAAGTGTGCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 4812
Db 3660 ATCACTTCAAGTGTGCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 3719
QY 4813 TATACGAGTGAACAATCCACCGTCTCAGAGCTTCAAGTGAATGTAATGTAATGTAATGTAATG 4872
Db 3720 TATACGAGTGAACAATCCACCGTCTCAGAGCTTCAAGTGAATGTAATGTAATGTAATGTAATG 3779
QY 4873 ATCCGATCAAGGAG 4932
Db 3780 TTCAAGTGAACAAG 3839
QY 4933 CTGAGAGTGAAGTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 4992
Db 3840 TTGAGAGAGAGTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 3899
QY 4993 GAGGAG 5052
Db 3900 GAGGAG 3959
QY 5053 AAGGAG 5112
Db 3960 AAGGAG 4019
QY 5113 CGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5172
Db 4020 CGGGGTGAGTCTCTCCCTTCAAGTCAATGATGATGATGATGATGATGATGATGATGATGAGGCA 4079

Dp	530	CAAGTTG3GTGAGCCTCACACCGGAAGAATTAACCTTTAGTGTCAATGAGCTCG	589
Qy	501	CTGAGCTTTGAGAGTATGCTCTCCTTAAGAGAACCTGGATGAGAACCTTCATGCGAC	560
Dp	590	GACCATGTCTGGTGTGATTAATGAGGTTCCAAAATGCGCAAAATPAACCTTTGTGTGGAC	649
Qy	561	GCGAGTGTAGGGAAGCCCGAGTATTTTCCACCATCTCCAGCCGGAACCTGACCAAGA	620
Dp	650	TCCATTATGATGAATAATCGAATATTTTCCACCTCTCCAGCCGGAACCTAGGAGAA	709
Qy	621	CTGTGAGGGGATGAGATGTTGCGGTGACCTTCATGATGAGATGTGTGGCCCTGCATGT	680
Dp	710	TGAAGAAAATGCTGAATATTTGGCTTTGTCTATCAGAGATGAAATTTGTGTCTTCAACT	769
Qy	681	TAAGATCCCTTGGGACACCTTCACCATCATCCCTGACTTTGATATCTATGTCTATGG	740
Dp	770	CAAGATCCCTCGATACCTTGTCAAAATTCCTCAATTTGATATCTATCTAGCTGTAGCAG	829
Qy	741	TTTTCACAGTGGCACTTTGTCTACTTTTTCACCTCCCAACCTGAGATGATGTCTCCAC	800
Dp	830	TTTTCACAGATGAGCACTTTGTGTACTATTAAACCTTCACCTGAGAACCCATTTAACTC	889
Qy	801	AGGCTCCACCAACGAAGAGCAGGTGTATACATCCACCTCTGTAGGCTTTGCAAGAGAA	860
Dp	890	ACCTGACTCAAGGGGAGCAGTTTATTCCTCCAAATGATCCGGCTTTGTGTAGAGGA	949
Qy	861	CACAGCTTCACCTCCTATGTAGAGGTCCCATTTGGCTGTAGACCGCAGTGGGATGAGTA	920
Dp	950	TCCCAAGTTCTACTCTTAAGTAAATTTCCATTTGGTGTGATGAAGAATGGGGTGAATA	1009
Qy	921	CCGCTGTGCAAGGCTGCCATCTGTCCAAAGCGGGGGCCGTGTGGCAGACCTTTGG	980
Dp	1010	TAGGCTGATACAGATGCTTATCTGAGAGCAACCGAGGAAACGCTTGGCTTAAGAGCTAG	1069
Qy	981	AGTCCATTCAGATGATGACCTGCTCTTCAACGCTTTCTCCAAAGGCGCAAGCGGAAAT	1040
Dp	1070	CATCTGTAAAGGAAGACATTTCTTTCAAGCTTTCTCAAGAGCAAGAAACCGAGT	1129
Qy	1041	GAAATCCCTGATGAGTGGGCCCTGTGTGATCTTATCTTGAAGCAATTAATGACCGAT	1100
Dp	1130	CAAAACCCCAAGAGTGTCTGTCTGTGCTCTTCACTTGAAGAAATCAAGATPAATAT	1189
Qy	1101	TAAAGAGCGCTGACAGTCTTTTACCGGGGGAAGGACACCTGAGACCTGGCTGTGCTAA	1160
Dp	1190	TAAAGCGGCATTTAGTCTTGTCTAAGAGGGGAATGGGAATCTCTCTGCTTGGCTGT	1249
Qy	1161	GGTGAAGACATCCCTGTGAGAGTGGCCTTTAACCATTTAGAGATTAATCTTCTGTGGCT	1220
Dp	1250	AAACAAAGGAGCTGGAGTGTGATCAATTCACTGTCAAAATTAAGACAAATTTCTGGGGGA	1309
Qy	1221	GGAATGAATGCTCCCTGTGGAGTGTCCGACATGTCGTGGAATTCGCCGTCTTACAGGA	1280
Dp	1310	GGAATTTAACAAGCGGCTTGGGGGGAAGTCAACATGAAGATGACACCGTGTCTTGGGA	1369
Qy	1281	GGAACAGGACCGCATGACGTCTGTATATGCAATATGTCTAACAAACCACTCTGTGGCCT	1340
Dp	1370	TAAAGAAAGCGGAATGAACATGATGACCGCTTAATATTCGTGGGGGAACAATGTGTGTT	1429
Qy	1341	TGTGGGACCAAAAGTGGCAAGCTGAAGAAAGATCCGGGTGATGGAACCAAGGGGCAACG	1400
Dp	1430	TGTGTGAACCGGAGTGGCCGTGTGAATAAATCTGTGTGGAACCTGTCTGCTCTCTC	1489
Qy	1401	CTCTC-----CAGTATGAGAACGCTGACAGTGTGTGAACCCCGGCCACAGTCTCCGGGA	1451
Dp	1490	TCACTTGTGTGAACAGTATGAGAAATGTATGTGTTCATGAGGGGCAATGTATCTTGAAGGA	1549
Qy	1462	TAAAGCCTTCTCCAAAGAACAGACCACTTACATCATGTCAAGAGGACGCTCAACAG	1511
Dp	1550	CTTGTACTAGAGCCCGACCGGCAATATATCTACGCTATGACAGAAAGAGATTAATCTG	1609
Qy	1512	AGTCCCTGTGTGAGTCTCTGTGTGCAATATACAGAGTGGGGGAATGTGCTTGTGCTCAGGGA	1571
Dp	1610	TGTCCCGGTGAGAGTGTGTGAACAAATGAAGAGCTGTGAACCTGTCTTGTGCTTCAAGGGA	1669

QY	1572	CCCCACCTGTGAGCTGTGTGTGTGCTGCACAACAACGTGCACCCCGGAAGAGACGGTGTGACG	1631
Db	1670	CCCTCACTGCGCGCTGTGTGTGTGTGCTCCACAACTGTGTGTTCGAGAAAAGATTAATGCGACGG	1729
QY	1632	GTCCAAAGAGCCCCCGCAGGTTTGCTTCGGAGATGAGACGATGTGTCCGGCTGACGGCTCA	1691
Db	1730	AGCAGATGAGATTACAGGTTCACTTGTGACCAACCGTCAAGTGTGTCCACTCAAGTGCA	1789
QY	1692	TCCCAACAATATCTCCGTCTCTCAGTACAACGTGTCTGTGTCTCTGTGAGACGTACAATGT	1751
Db	1790	CCCCAAAAATATCTCCGTCACTGTGTCCAGAGTTCCATGGTTTAAACCTGTGAATGT	1849
QY	1792	CCCGAGAGTGTCAGCTGTGCGTCACAATGACACTTGTGAGACCTGTCAAGATNGATGGAGCT	1811
Db	1890	CCCTGACCTGTGTCTGTGGGTGAACTGTCTCTTTGAGSATTTACTGATGATGAGAGACG	1909
QY	1812	GGTGTGGGCAATCAATCCAGTGTACTCTCCCTGACGACCAAGAGGTGCCCGGATCAT	1871
Db	1910	GATCTGATGTGAAGATCTACTGCAATCTCCCTCTGTCAAGAAATTTATCTTCACTAC	1969
QY	1872	CACAGAGATGGGAGCCACCATGTGTGACAGCTTCAGCTCAATATCAAGAGACCCGCAAT	1931
Db	1970	AAGGGGCAATGAGACAAAGAGGTTGTAAACTGTACTTAAATCCAAAGAAACTGGGAA	2029
QY	1932	GACCTTGGCCAGACCAAGCTTTGTCTCTCAATTCGACGGGTCCACAATTCGTGCTGTCTC	1991
Db	2030	GAACTTTGCAAGTGTGATTTTGTCTTCTCAACTGTAGTGTCCATCAGTCATGACCTCTC	2089
QY	1992	CTGTGTGAGAGTGTCAATACGCTGTCCACTGTGTAAATACCGGATCTCTGCAACCATGA	2051
Db	2090	CTGTGTAAATGATTCCTTCCCTGTCCACTGTGTGAATATAGACACTGTGTCAACCCCAAA	2149
QY	2052	CCCCAAGACCTGTCTCTTCCAGAGAGGCCGAGTGAACCTGCCGAGGACCTGCCCCCACT	2111
Db	2150	TGCAACCGACCTGTCTCTTCAAGAGGSCCGAGTGAATGTGTGAGGACCTGTGCCCAAT	2209
QY	2112	GCTGCGAGTGAACAAGATCTGTGTGCCGTGTGAGGTGATCAAGCTTATCAAGCTGAAGCC	2171
Db	2210	TCTGCTCTTCCCTCCAGATCTACATCTCTGTGTGGGTGTGAAAGCCATCACTACGACTGC	2269
QY	2172	CAAAACCTCCCCCAGGCCAGGTGTGGGACGCTGTGAGATGATTCCTCAACATTTCA	2231
Db	2270	CAAGAACTCCCGGAGGCCCAAGTCCGCTCAGCGCAATTAGAGTATTTTCCACATCCC	2329
QY	2232	GGGCAAGCAGACGAGTGTGCCCGCTGTGCTTCAACGTCTCCAGGTTACAGTGTGCCGAA	2291
Db	2330	CGGAGAGTGTACTGGGTCACTGTGTTCGCTTCACAGTACACAGATCCAGTGTCCGAA	2389
QY	2292	CACCTCTTATCTCATGAAGGATGTGATATCAACACTGTGCCGTGTGAGTTGAACATGCT	2351
Db	2390	CACCTCGTATACATTAAGGCAATGATATCAGTGACTTACCTGTCAATTTATCTGTGGT	2449
QY	2352	GTGAATGTGGACATTTCAACTTGAACAACCCGACTCAATATAAGTTCACCTTCAACATGT	2411
Db	2450	CTGAATGTGGACATTGTGTCAATGACAAACCTCAAAATTCATACCTCATCTGTACAAAGT	2509
QY	2412	TGAAGCAATGCTGTGAGAGCTGTGGGCTGTGTGCTCAAGGCTGACCCAGACTTGTGATGTG	2471
Db	2510	CTCGGCTCTTAAGGAAAGCTGTGCGGACTTTGCTTAAATCTGACCGTGTGATTTGTGTGTG	2569
QY	2472	CTGTGTGACGAGGCCCAAGGCCAGTGCACCTGTGCGGACACATGCGCTGCCAGAGAGCA	2531
Db	2570	CTGTGTGTGCTCAGTGAAGAAAGTGTCACTGTGTGTGCAABAATGTGTCCACCTTGAANAACC	2629
QY	2532	GTGTGTGAGCTGTCTGTGTGTCAAAAGCAAGTGTCAAAACCCCGCATCACAGATTAAT	2591
Db	2630	TTGTGATGTATCTCAATACCGCAATAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2689
QY	2592	CCCGGTACAGGCCCGCGGAAAGGGGACCAAGAGTCACTATCCGAGGGGAAACTGTGG	2651
Db	2690	TTCAGAGACAGGCCCCAGGCAAGGTGTGAACCCGACTACATCAACAGAGAAATCTGTGG	2749

QY 2652 CCTGGAATTTCCGACATGCGCTCCCATGTCAGAGTTGCTGGCGTGGAGTGACGCTTT 2711
Db 2750 CTTGAGGTTGAGAGACATTCGCTTTGAGTGAAGGTTGGGACGTCATGTGTGCTCCGT 2809
QY 2712 AGTGGATGTTAATCCCTGACAAACAGTTCGTGTGTGAGATGGGGAGGCC---AAGCC 2768
Db 2810 CGAGAGGATATATACGCGACAGACAGATCGTTGTGAGATTAAACATGCCGAGAAC 2869
QY 2769 CAGCAGCATGACAGGCTTCGTGAGATTCGCGTGTGTGTGTGCGGCTGAATTCATGGC 2828
Db 2870 CCGTGTCCACGAGCTCAAGTGAAGGTCTGTGTAAAGACTGCTCTCAAGACTACAGGC 2929
QY 2829 CCGTCTCCACAGCTCATATTACTGACACTGACTCTCTCAATCTGAAGCCAGCG 2888
Db 2930 TATTCTCCAAAGTCATTCACCTTGTGTACCTCTCTTAAACGGGTGACACTTCCG 2989
QY 2889 GGGGCCCATGTCGAGGAGGCCCAAGTGAACATCAAGGACCAACTGAATGCCGAG 2948
Db 2990 AGGTCTTGTGTGGGACCTGGATTCATTTGAAGAACTATCGAAGCTGGCAG 3049
QY 2949 CAACGTGTGTGATGTTGGAAGAGAGCGCTGTCTTCCACAGCGATCTCATCTTA 3008
Db 3050 TGAATGTTCTGTGCTATTTGGGAGGCTCCCTGCAATGTTTTCATGGCAATCGGAAAG 3109
QY 3009 CATGTCTGACACACACATCTCAGATGAGTGTAGATGAAGTGTGAGTCAGT 3068
Db 3110 GATCCGCTTAAAGCTCCGAGGAGCAAGCAACGGGTAAAGCCGAATACAAATCTCAT 3169
QY 3069 GAGACGGGCCAAGAT---CCACAGGACCTGTGCTTTCAGTATGTGGAAGCCCAACAT 3125
Db 3170 CAACCGGCGACATGAACATCTTAAGTACATTTACATTCACCGAAGATCAACCT 3229
QY 3126 CGTGGATTTAGCCCAAGATGAGCATTTCTAGTGAACACACCCATCGCGTATGGG 3185
Db 3230 GCAAAAGATAGACCCAGAGTGGAGCATCCGCAAGTGGGAGAACCTCTGATATGTAACG 3289
QY 3186 GACCCACCTGACCTCAACAGAACCCCAAGTCCGTCAGACATGGAGGAGAGAGAC 3245
Db 3290 GATGATCTGSCAATTTAAAGAACCCCAATTAAGCTAATATAGAGATGTGAGAA 3349
QY 3246 CATCATATCTGTGAGGTTCTGAACGCTAATGATGACCTGTGAGCGCGCCCTGCG 3305
Db 3350 AGAAATATACGACTCTTTACATATGACACACATGTGTGCTTGGACCTCTGCTGA 3409
QY 3306 TCTGGTCTGACCAACAGTCAACCTGACCGAGAGCCGAGAGTTGGCTTCACT 3365
Db 3410 CAATCCTTGAAGAGCCCTCCGAGATGGGAGCGTCTGTATGATATAGCTTCAAT 3469
QY 3366 GAGACATGTCAGTCTCTGCTCATCTCAACAGACCACTTACCTTATCCAACCC 3425
Db 3470 GGAATATGTTCACTCCCTCTCTTATGTGAACATACACATCTCTTATTAACCGGATCC 3529
QY 3426 GATGTTGAGGCTTTGATCCCTCAGAAATCCCTGAGCTCAAGCTGGACGCGCATAT 3485
Db 3530 TGTGTTGAACCCCTGACGCACTGTGAACCTGTGAGTGAAGCCCAAGTCTCTTAT 3589
QY 3486 CTTAAAGGCAAGAACCTGATCCCGCTGTGCTGGGAGCACTGAAAGCTGAATAC 3545
Db 3590 AATCAAGGCTGGAATTCATTCAGACACCCCAAGG---AATCTCAGGCTTAATTAAC 3646
QY 3546 TGTGCTGTTGGGAGAGCCGTGACACCGTCAAGTGTCAAGTGTCAAGTCTTGA 3605
Db 3647 AGTGTGATTTGTGACACTCCGTGTCTTACTGTCTCGAGACCCAGCTCTGTGGA 3706
QY 3606 GTCCCCCACTCATCGGACGACCAAGTGTGCGCGCTGTGCTGTGAGATTC 3665
Db 3707 GTCTCCAACTTACGTGACAGCATAAAGTTACTATTAAGCCGAGGAGTTGATATTC 3766
QY 3666 CCGCGGATGATGATATGCCCCGAGACAGCCGCTCAAGCTGCCCCGATGTCAGAT 3725
Db 3767 ACCGGGACACTGAGATTTACTCGGACAGCCGTGACCTTGTCTGCAATTAATGGAT 3826
QY 3726 CGAGTGTGCGGCGCTCTCATCATTTGATCGTGGCGGTGCTCATTTGCTTAACG 3785

Db 3827 TGAAGAGAGAGGGGTTTGTGCTGATCATTAATATCATGTGTCTATTCCTAACAGAG 3886
QY 3786 CAAGTCCGCGAAGATGACCTCAAGCTGAAGCGGCTCAGATGAGATGAGCAACCTGA 3845
Db 3887 GAATGTGGGATGCGGACCGGACCTTGAAACGACTGAGCTGAATGAGCAATCTGGA 3946
QY 3846 GTCCGTGTGCTCTGAGTGAAGAAAGCTTTGCGAGCTGACAGGACATTCATGA 3905
Db 3947 GTCCGTGTGCTCTGAGTGAAGAAAGCTTTGCGAGCTGAGCTGACATTCACGA 4006
QY 3906 GCTGACCAAGTACCTGATGAGCGGAGATTCGTTCTGTAAGCTATTAACCAT 3965
Db 4007 GCTGACCAAGCTGAGCGGCTGGAATTCCTTCTGAGATGAGCTGATTAACCAT 4066
QY 3966 GCGGTGCTGTTCCAGGAATTAAGACACCTGCTCTCGGAGCTTGAAGTCCCGG 4025
Db 4067 GCGAGTCTTTTCCCGGATGAAGATCATCTGTCTGAAGAAATGAGGATCAAGC 4126
QY 4026 CTACCGGACGAGCGTGTGAAGAGGCTGAAGCTCTGCGCCAGCTCATCAACAA 4085
Db 4127 CAAC-----GTGAAGAAATCCCTGACGCTTTTGGGAGCTCTCTCAACAAAG 4174
QY 4086 GGTGTTCTGCTGTCTTATCCGACGCTTGAAGTCCCAAGCTTCTTCAATGCGCA 4145
Db 4175 GCACTTCTGTTGACCTTCAATTGAGCGCTGAGAGCCCAAGAGCTTCTCAATGAAGA 4234
QY 4146 CCGTGAACGTTGCTCACTCATATGACCGTCTGACAGCAAGCTGAGTACGAC 4205
Db 4235 CAGGGTAAAGTGTGCTGCTCATATGACGCTGACAGGAGATGAGAAATACGCA 4294
QY 4206 TGAATGCTGAAGACGCTGCTGCGGACCTGATTTGAACAACTGAGAGCAAGCA 4265
Db 4295 CGGCTTCTCAAGACGCTTCTCTCGACCTTATGAAGAACTTTAAAGCAAAACCA 4354
QY 4266 CCTAAGCTGCTGCTCAAGAGAACTGATCAAGTGTGAGAGATGCTGACCAATGGTT 4325
Db 4355 CCTAAGCTGCTGCTGAGAGAGCAAGTGTGAGTGTGAGAGATGCTGACCAATGGTT 4414
QY 4326 TACTTCTCTCTCTCAAGTCTCTCAAGATGTGCGGAGAGGCGCTCTTCCCTGTT 4385
Db 4415 TACGTTCTCTCTCAAGTCTCTTAAGAAATGTGCGGAGGCGCTCTTCAATGTGCA 4474
QY 4386 CTGTCATCAAGCAGATGAGAGAGGCGCCATTAAGCGCATCAAGGAGGCGCG 4445
Db 4475 CTGTCATCAAGCAGATGAGAGAGGCGCTTATTAAGCGCATCAAGGAGGCGCG 4534
QY 4446 CTACTCTTGAAGGAGCAAGCTCATCCGACAGATTAAGCTAAGAAACCTG----- 4500
Db 4535 CTACTCACTGAGCGAGCAAGCTGATCTGACAGATGACATAAGACCTGAATCC 4594
QY 4501 -----GTCTGAG 4508
Db 4595 TTGTGTGATGACTTTGGGCTGAGAGATGATCATGTTGCGCTCGCTCAAAACCTGAA 4654
QY 4509 CTGTGTGAGCCAGCAATGCAACAGCCCGAGGCTCCAGTAAGATCTCAACTGTA 4568
Db 4655 CTGTGTGAGCCAGCAATGCAATGCAAGATGCTCCAGTAAGATCTCAACTGTA 4714
QY 4569 CACATCACTCAAGTCAAGAGAGATTTCTGATGCACTTCAAGATGTGCTTGTCTC 4628
Db 4715 CACATCACTCAAGTCAAGAGAGATTTCTGATGCACTTCAAGATGTGCTTGTCTC 4774
QY 4629 CCAAGGCGCAAGTGTGAGATGATCTGAGTGTGAGCAACAGAGAGTGGGCAAGAT 4688
Db 4775 CAGAGCAACAAAGGAGAGACATGACCTGAGTGTGCGAGGATGATGAGACCAT 4834
QY 4689 GATCTTGAAGTGAAGCATCACCAAGATTTGAATGATTTGAAGCATGGAAC 4748
Db 4835 TATCTGCAAGATGAGTGTGACCAACCAATTAAGCAATGACCTGGAAGCACTAATAC 4894
QY 4749 ACTGGCCACTCAAGTGTGAGATGTTCTGATGAGATTTAGTGTCAAGCAGGTGAC 4808

Db 4895 ATTGGCATTATTCAGTGTGACAGATGATGATCTTCACTGCTGATCCCAACAGACATC 4954
 Oy 4809 AGCTTATACGCGAGTGAACAACTCCACCGTCTCCAGACCTCGCAAGTAAATATGAAA 4868
 Db 4855 TGCTTACAAATCTCAACATCATATCACTTAAACAAAG---TCTGTAGAGATATGAGAG 5011
 Oy 4869 CATATCCGGTACACGGGACGCCCCGACGCTCCGCTACAGACACCTATATCATCTCC 4928
 Db 5012 TATCTTACGACGCGATCAACCCAGACACCTGAGCTCCGACCCCATATGATCACTCC 5071
 Oy 4929 TGACCTGAGAGTGTGAGTCAAGATGTGACCTTATGAGAACCCAGACACGAGACCA 4988
 Db 5072 TGACCTGAGAGTGTGAGTCAAGATGTGACCTTATGAGAACCCAGACACGAGACCA 5131
 Oy 4989 GAAAGAGAGGAGACCGGGGAGACAGATGTGTGAAATCTACTACCGGACTCCTGCGC 5048
 Db 5132 ACGGAGAGGAGTACAGAGATAGAGATGTGTGAGATTTTACCTTACCTGCTGCTGTTGCG 5191
 Oy 5049 CACTAAGGACACTGACAGAAATTTGTGATGACCTCTTTGAGACCATCTTACAGACGCG 5108
 Db 5192 CACAAAGGAACTCTCCAGAAATTTGTGATGATGTTGAGACGATCTTACAGACGCG 5251
 Oy 5109 ACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5168
 Db 5252 ACATCTGTGAGAGCGCGCTCCTCTCCTGATTTAAATATGATTTGATTTCTTGAAGAGCA 5311
 Oy 5169 GCGTATTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5228
 Db 5312 GCGGAGACAGATCAGATCAGATCAGATCAGATCAGATCAGATCAGATCAGATCAGATCAGAT 5371
 Oy 5229 GCGCTGATTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5288
 Db 5372 GCGCTGATTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5431
 Oy 5289 GAAACAGATCAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5348
 Db 5432 GAAACAGATCAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5491
 Oy 5349 CACCTGAGAGACCGGCTGGGACAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5408
 Db 5492 CACCTGAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5551
 Oy 5409 CATCCCAAGCTCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5468
 Db 5552 CATCCCAAGCTCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5611
 Oy 5469 CATCCCAAGCTCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5528
 Db 5612 GATGAGGACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5671
 Oy 5529 GTTCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5588
 Db 5672 GTTCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5731
 Oy 5589 GATCCCTGAGACCTGAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5648
 Db 5732 GATTTTAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5791
 Oy 5649 AGAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5688
 Db 5792 GGAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5831

RESULT 15
 Db 5831 GGAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5831
 Db 5792 GGAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5831

DEFINITION Mus musculus mRNA for plexin 3, complete cds.
 LOCUS D86950 5982 bp mRNA linear ROD 31-MAR-2000
 VERSION D86950.1 GI:1655433
 KEYWORDS plexin 3.
 SOURCE Mus musculus (house mouse)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 2 (sites)
 Shimizu, M., Murakami, Y., Suto, F. and Fujisawa, H.
 Determination of cell adhesion sites of neuropilin-1
 J. Cell Biol. 148 (6), 1283-1293 (2000)
 20191894
 10725340
 3 (bases 1 to 5982)
 Fujisawa, H.
 Direct Submission
 Submitted (07-AUG-1996) Hajime Fujisawa, Nagoya University Graduate
 School of Science, Division of Biological Science, Furo-cho,
 Chikusa-ku, Nagoya, Aichi 464-01, Japan
 (E-mail: fujisawa@bio.nagoya-u.ac.jp, Tel:052-789-2978,
 Fax:052-789-2979)

FEATURES
 SOURCE
 CDS
 1..5982
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /tissue_type="brain"
 /dev_stage="E17.5"
 144..5762
 /codon_start=1
 /product="plexin 3"
 /protein_id="BA13190.1"
 /db_xref="GI:1655434"
 /translation="MPTVCLPLFTTGGCTGSSRPRTFVTDITTLIAVHTG
 EYFVGVNVEFLAPNLTLELRAVTPTEEDNARCPEPSMVCGRILVPAVNVKLL
 IDYAPRLVACDSIWOGLCOPLDLPLKSEPHRKSHYVQDEPSSQIKIPDTS
 OGSFRLVGTADGKSEYFPLTSRKLIDDESGMSEFLVYQDFVSSQIKIPDTS
 LYAPFIYIYGFVNASFYFLTLDTQDTLDTAGAEFTSKIVKACGDSSEFY
 VEPFGCGRWGVRLVQSAHAKKGLLAQALGVPADEVDLFTFSGGGRNANP
 OTICFTLISNAHRRIOSCYRGSEDTLPLINLEICINPLOTIONFGVIL
 NPLGSLATIBELPLADSTDMASVAAYTHHSHVYIGRSGNKKYRDGSDAQ
 LRETVSVAGSTPLDLPSPDHRTYLSKQVSOQVEFCEQTLSCACLGSDPH
 CMCVYQHRCCREKACPGASAPHGAEELSCIOVRPNVSVTSSGVQTLVMRV
 PDLVSVGSCSFEVTESEALPLPSGELCPSPISLOELTLRGGAHTYRLQLSME
 TGVRPAGVDVFPYNCALQS CMS CVGSPYPCMCYRYVCTSHPHSCFQGRVHSP
 GCEILIPQDLIPVGNOPFLRAKNIPOSGGKNCVCVVRQGRHVPVAFNS
 SSVQCNMAEYFESGDEPDELDFSVWMDGSPEDIKPSFALLIKCAQRPSCGL
 KADPNCNCGMSHRCOLRAICPAKSNMWHPSCKGRCSHPRIITOHPLTGKRG
 TRVTVGENLGLTSREVGIRVAVACNSIPREYSAERIVCEMESLYSPSPPAEL
 CVGDCADRTQSQQLISFVTFPDRVSPSPKSGGRTLTSGSLDAGSRVTLIR
 DECCQFRRDAALIVCISPVSTLGPSPPTILADHANIISVGIYITQSPVTLIR
 PTWSIINGSTSIIVSGTHLITVOERVARAKGLETNTTCOVINDTALCPAPIFLG
 HPQPRQSGHDEFEFLDHYQARSLSRSGFTYPPDPSPPLGSPGVLDKPSHVV
 LKGNLIPAAASSRLANTYVLIQGPCLATYSDROLDSQSGQCPVMDLVGSLER
 MGLHITIDRALTLTPAMVGAAGGGLLITVLLVYKRTQDADRTLRLOMD
 NLESRALECKEFAELQDINELTNHMDGOIPLDYRTAVAVLPGIIGAHVLE
 LDPNVEKALTELPGOLLSRAFLITFTHLTAQSSREMRDGTASTYVLAISRLD
 YATGLKQDLADLIRKLSKNEPKLIRTESVAKRLTWMTFTLHKRLEKQAGEP
 LPLTYCALKQNEKRPIDALITGEARYSISEKULRQOQDYRTTLHCPSPESDAQV
 PKVINCOSITQAKDKLIDTVYKGIPIYQRPKADMDLEWQGNARILIDDEDITTK
 IECDMKRVNSLVAHVQVTDGSLVALVAPKQVSAVYNNANSTFTFRSLRSLRAASPD
 SLRSRAFMLTPOEAGTKLMHLVRNHDTDRBDGRSKVSELYTFLTLTKGTLOK
 FVDDLEFVETSTARRGSAALPLAIKMPFDLBOADROIIDPVATHTKSCLEPLRE
 VVIVKQPOFVEDIHNSITDACLSVAQDTFSDSCSHRGLKQSPSKLLAYADIPN
 YKSWERYRDIARKWASTSDODMAVYVQSGRLHNDPNTVLSALSELYFTYTKRQEI
 LNSLDKASCRGHKRLQKPEQILITVSSSS"

ORIGIN
 Query Match 40.6%; Score 2312.8; DB 10; Length 5982;
 Best Local Similarity 64.7%; Pred. No. 0;
 Matches 3592; Conservative 1917; Indels 39; Gaps 9;

Db 2399 TGAAGTATGATGTTGGTGAACCGAAGTGAATCTTCTGTGGTTGGATGGAATTT 2458
QY 2367 CAACATTTGCAACCCAGCTCAGAAATAAAGTTCACTCTAACAAGTGTGAGGCATGCGTGA 2426
Db 2459 CCCCATTTGATAAGCTCCCTAGCTCCGAGCCCTTCTTTAACAAGTGTGAGGCATGCGGCC 2518
QY 2427 GAGCTGCGGCTGTGCTCAAGCTGACCCAGACTGGCATGTGGCTGTGAGGCATGCGGCC 2486
Db 2519 TAGCTGTGCTCTGCTCTAAGCTGATCCCGATTCACTGTGTGCTGTGATCTAGA 2578
QY 2487 AAGCCATGTCACCTGTGCGCAGCACTGCTGCGCAGGAGCAAGTGTGAGCTGTG 2546
Db 2579 GCAAGTGTGCTGAGTGGAGCTCACTGCGCAGCTCCAAAAGTATGTCACCAAG 2638
QY 2547 TGGTGCCAAAAGCAAGTGAACAACCCCGCATCAAGAGATTAATCCCGGTGACAGGCC 2606
Db 2639 CAGAAAGGTGCGCCATGCAAGCATCCCGAATCAACCAATTCATCACTTAAGAGACC 2698
QY 2607 CCGGAAAGGGGACCAAGTCACTATCCGAGGGAGAACTGAGCTGTGAATTTGCGA 2666
Db 2699 CAGAGAGGTGACACCGGCTCACCATGTGAGTGAACCTGAGGCTCACTTCCGCTGA 2758
QY 2667 CATGCGCTCCCATGTCAGAGTGTGGGTGAGTGAAGCCCTTATGATGATGTTATCAT 2726
Db 2759 GGTGGCTTTGAG-----TAGCTGTGTAGCTTGAACCTCATCCACCAATATGT 2812
QY 2727 CCTGTGAAACAGATGTGTGTGAGATGAGGGA--GGCCAAAGCCCAAGCAGATGCAAG 2783
Db 2813 CAGTGTGAAAGATGTGTGTGAGATGAGGAAATCGTGTGCGCCAGCCCAACCACTGG 2872
QY 2784 CTTGCGTGAAGATGT 2843
Db 2873 GCGTGTGAGCTCTGT 2932
QY 2844 CTATTAATTCATGACATGATCTCTCATGATGTGAAGCCCAAGCGGGGGCCCATGTCCGG 2903
Db 2933 CTACAGCTTTGTGACCCCTTACATTTGACCGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2992
QY 2904 AAGGACCCAAAGTGAACCATCAAGGACCAACCTGAAATCCCGAAAGCAACGTGTGTGTGT 2963
Db 2993 AAGGACCTGGCTTACATCTCTGAAATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3052
QY 2964 GTTTGAAAGCAAGCCCTGTCTTCTTCAAGGCGATCTCATCTTCACTTGTGTGCAAC 3023
Db 3053 TATAGAGATGCGAGTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3112
QY 3024 CACATCTCAGATGAGT 3080
Db 3113 ACTGT 3172
QY 3081 GAT--CCACAGAGACCTGT 3137
Db 3173 CATCTCCAACTGT 3232
QY 3138 GCCAGATGAGATGT 3197
Db 3233 GCCCACTGT 3292
QY 3198 CCTCATAGCAACCCCGAGATCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3257
Db 3293 GACAGTTCAAGAACCCCGT 3352
QY 3258 TGAAGTTCGAAAGCTGT 3317
Db 3353 CCAAGT 3412
QY 3318 CCAAGT 3377
Db 3413 TCAAGT 3472
QY 3378 GTTCCCTGT 3437
Db 3473 GGAAGCCCGCTGT 3532

QY 3438 CTTGTGCTCTGAGAAATCTGTGAGCTCAAGCTGTGACGCCCATCATCTTAAAGGCA 3497
Db 3533 ACTTGGGCTCTGT 3592
QY 3498 GAACTGTATCCCGCTGT 3557
Db 3593 GAACTGTATCCCGCTGT 3649
QY 3558 GAGAAAGCTGT 3617
Db 3650 AAGCAAGCAATGT 3709
QY 3618 CATGCGGAGCAAAAGT 3677
Db 3710 GACAGGCGGAGCTGT 3769
QY 3678 GTACATTTGCGGAGAGCTGT 3737
Db 3770 GCAATTCATCTGT 3829
QY 3738 CAGCTCTCTCATATTTTATGT 3797
Db 3830 CAGGCTCTGT 3889
QY 3798 AAGT 3857
Db 3890 TGCAGACCGACACTTAAAGCGGCTTCAAGTCAATGTGTGTGTGTGTGTGTGTGTGTGT 3949
QY 3858 CTTGAGT 3917
Db 3950 CTTGAGT 4009
QY 3918 CTTGAGT 3977
Db 4010 CATGATGT 4069
QY 3978 CCAAGAAATTAAGAACCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4037
Db 4070 CCGT 4120
QY 4038 GCGT 4097
Db 4121 --TGTGAGAAAGCCCTGT 4177
QY 4098 GTCTTCAATCCGACGCTGT 4157
Db 4178 CACCTTTATCCAACTTTGAGAGGCCCAAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4237
QY 4158 GGCCTCATCTCATATGACCTGT 4217
Db 4238 GGCCTCATCTCATATGACCTGT 4297
QY 4218 GCAGCTGT 4277
Db 4298 GCAATCTGT 4357
QY 4278 GCTCAGAGAGCTGT 4337
Db 4358 ATTGCGAG 4417
QY 4338 CTACAGT 4397
Db 4418 GCATTAAGT 4477
QY 4398 GCAGAGATGT 4457
Db 4478 GCAGAGATGT 4537
QY 4458 CGAGGACAGT 4517
Db 4538 TGAAGCAAGCTGT 4597

Mon Feb 23 09:34:02 2004

us-09-964-956-12.rge

Page 46

Job	Time	Search completed:	February 19, 2004, 20:48:07
Db	5675	CTCGCTGGAGCCAGATGCTCTTGTGGAGGACAAAGCTTGACAGAAAGCTGAGCAGAT	5734
Qy	5658	CATTAACCTCATGAGCTTGAACAAGCTGA	5685
Db	5735	CATCAACCTGAGTGTCCAGCAGAGCTGA	5762

[illegible]

QY	132	AGAAACA	CTCTCTGGCCCTTTGTGGGAC	CAAAAGTGG	AAAGCTGA	AGAAAGATCCGGGTG	1380
Db	1321	AAAGACA	CTCTCTGGCCCTTTGTGGGAC	CAAAAGTGG	CAAGCTGA	AGAAAGATCCGGGTG	1380
QY	1381	GATGAGC	CCAGGGGCAACGCCCTCCAGAT	ATGAGACGGTGC	AGGTGGTGGACCTCCGGCCA		1440
Db	1381	GATGAGC	CCAGGGGCAACGCCCTCCAGAT	ATGAGACGGTGC	AGGTGGTGGACCTCCGGCCA		1440
QY	1441	GTCCCTCCGGGATATG	AGCTCTTCTCCAAAGGAC	CAAGAGAACTCTA	CAATCACTGCAGAGAG	1500	
Db	1441	GTCCCTCCGGGATATG	AGCTCTTCTCCAAAGGAC	CAAGAGAACTCTA	CAATCACTGCAGAGAG	1500	
QY	1501	CAGCTCA	CCAGAGTCCCTGTGGAGTCC	TGTGTCA	GTATCAAGCTGCGGCA	ATGCTT	1560
Db	1501	CAGCTCA	CCAGAGTCCCTGTGGAGTCC	TGTGTCA	GTATCAAGCTGCGGCA	ATGCTT	1560
QY	1561	GGCTCAG	GGGCAACCCCACTGGTGTG	ATGTGTGCA	CAACATGTCAC	CCGGAAGAG	1620
Db	1561	GGCTCAG	GGGCAACCCCACTGGTGTG	ATGTGTGCA	CAACATGTCAC	CCGGAAGAG	1620
QY	1621	CGGTGAG	GGGCTCAGAGACCCCGAG	TTTTCCTCCGAGAT	ATGAGCAATGCTGTCCG	1680	
Db	1621	CGGTGAG	GGGCTCAGAGACCCCGAG	TTTTCCTCCGAGAT	ATGAGCAATGCTGTCCG	1680	
QY	1681	CTAGCAGT	CCATCCCAACATATCCCGAT	CTCAGTACA	CAATGCTGTGCTGTCCGAG	1740	
Db	1681	CTAGCAGT	CCATCCCAACATATCCCGAT	CTCAGTACA	CAATGCTGTGCTGTCCGAG	1740	
QY	1741	ACGTACA	TGTCCCGAGCTGTCA	GTGGCTCAATGCA	CTTCTTGAAGACCTGTCA	AG	1800
Db	1741	ACGTACA	TGTCCCGAGCTGTCA	GTGGCTCAATGCA	CTTCTTGAAGACCTGTCA	AG	1800
QY	1801	ATGATAT	GAGCTGGTGTGTGTGAG	CAATGATATGAT	CTTCTTGAAGACCTGTCA	AG	1860
Db	1801	ATGATAT	GAGCTGGTGTGTGTGAG	CAATGATATGAT	CTTCTTGAAGACCTGTCA	AG	1860
QY	1861	CCCCGAT	CATCA	CAGAGATGGGACCA	CAATGTCTCA	AGTTTCACTCAATCA	1920
Db	1861	CCCCGAT	CATCA	CAGAGATGGGACCA	CAATGTCTCA	AGTTTCACTCAATCA	1920
QY	1921	GAGAC	GGGCA	TGACCTTGC	CAGACCA	AGTTTGTCTTCA	1980
Db	1921	GAGAC	GGGCA	TGACCTTGC	CAGACCA	AGTTTGTCTTCA	1980
QY	1981	TGCTGCT	GTCTGTGTGAGTGTCA	TATACCGCTGTCA	ATGATATGAT	CTTCTTGAAGACCTGTCA	2040
Db	1981	TGCTGCT	GTCTGTGTGAGTGTCA	TATACCGCTGTCA	ATGATATGAT	CTTCTTGAAGACCTGTCA	2040
QY	2041	TGACCA	ATGACCCCA	AGCTGTCTTGC	CAAGAGACCTGTCTTGC	CAAGAGACCTGTCTTGC	2100
Db	2041	TGACCA	ATGACCCCA	AGCTGTCTTGC	CAAGAGACCTGTCTTGC	CAAGAGACCTGTCTTGC	2100
QY	2101	TGCCCCA	GTGTGTGTGAGTGTCA	TATACCGCTGTCA	ATGATATGAT	CTTCTTGAAGACCTGTCA	2160
Db	2101	TGCCCCA	GTGTGTGTGAGTGTCA	TATACCGCTGTCA	ATGATATGAT	CTTCTTGAAGACCTGTCA	2160
QY	2161	ACGCTGA	AGGCA	AGCTGTCTTGC	CAAGAGACCTGTCTTGC	CAAGAGACCTGTCTTGC	2220
Db	2161	ACGCTGA	AGGCA	AGCTGTCTTGC	CAAGAGACCTGTCTTGC	CAAGAGACCTGTCTTGC	2220
QY	2221	CTGA	CAATTCAGGGGCA	GGGAGAGATG	CGGCTGTCA	CAAGCTTCAAGCTTCA	2280
Db	2221	CTGA	CAATTCAGGGGCA	GGGAGAGATG	CGGCTGTCA	CAAGCTTCAAGCTTCA	2280
QY	2281	CAGTGC	CAAGACCTCTTAT	CTATGAGAGAT	ATGATATGAT	CTTCTTGAAGACCTGTCA	2340
Db	2281	CAGTGC	CAAGACCTCTTAT	CTATGAGAGAT	ATGATATGAT	CTTCTTGAAGACCTGTCA	2340
QY	2341	TTGAC	ATGTGTGTGAGTGTGAG	CAATGATATGAT	CTTCTTGAAGACCTGTCA	ATGATATGAT	2400
Db	2341	TTGAC	ATGTGTGTGAGTGTGAG	CAATGATATGAT	CTTCTTGAAGACCTGTCA	ATGATATGAT	2400

QY	2401	CTTACAACTGTGAGGCCATCTCGTGAAGAGCTGCGGGCTGTCTCAAGAGCTGACCCAGAC	2458
Db	2401	CTTACAACTGTGAGGCCATCTCGTGAAGAGCTGCGGGCTGTCTCAAGAGCTGACCCAGAC	2466
QY	2461	TTCCGATCTGGCTGGTGTCCAGGGGCCCAAGCCAGTGCACCTTGCGCCAGACATCTCCCTGACC	2520
Db	2461	TTCCGATCTGGCTGGTGTCCAGGGGCCCAAGCCAGTGCACCTTGCGCCAGACATCTCCCTGACC	2520
QY	2521	CAGGAGAGCCAGTGGCTGAGAGCTCTCTGTGTCTCCAAAAGCAAGTGCACAAACCCCGCATC	2580
Db	2521	CAGGAGAGCCAGTGGCTGAGAGCTCTCTGTGTCTCCAAAAGCAAGTGCACAAACCCCGCATC	2580
QY	2581	ACAGAGATATATCCCGGTGACAGGCCCCCCGGGAAAGGGGCACCAAGTGCATATCCGAGGG	2640
Db	2581	ACAGAGATATATCCCGGTGACAGGCCCCCCGGGAAAGGGGCACCAAGTGCATATCCGAGGG	2640
QY	2641	GAGAACCTGGGACCTGGAATTTCCGAGACATCGCTCCCATGTCAAGGTGTCTGGCGTAGAG	2700
Db	2641	GAGAACCTGGGACCTGGAATTTCCGAGACATCGCTCCCATGTCAAGGTGTCTGGCGTAGAG	2700
QY	2701	TGCAGCCCTTTAGTGGATGTTTATATCCCTGCAAAACAGATCTGTGTGAGATGAGGGAG	2760
Db	2701	TGCAGCCCTTTAGTGGATGTTTATATCCCTGCAAAACAGATCTGTGTGAGATGAGGGAG	2760
QY	2761	GCCAAAGCCAGCCAGCATAGCAGGCTTCGTGGAGATCTGCGTGGCTGTGTGTCCGACTGAA	2820
Db	2761	GCCAAAGCCAGCCAGCATAGCAGGCTTCGTGGAGATCTGCGTGGCTGTGTGTCCGACTGAA	2820
QY	2821	TTTCATGGCCCGGCTCTCTCAAGCTTATTACTTCATGACATGATCTCTCTCAGATCTGAG	2880
Db	2821	TTTCATGGCCCGGCTCTCTCAAGCTTATTACTTCATGACATGATCTCTCTCAGATCTGAG	2880
QY	2881	CCCAAGCCGGGGGCCCATATGCTCCGAGGGAGCCCAAGTGCACATCAGAGCACCAACTGAAAT	2940
Db	2881	CCCAAGCCGGGGGCCCATATGCTCCGAGGGAGCCCAAGTGCACATCAGAGCACCAACTGAAAT	2940
QY	2941	GCCGGAAGCAACGTGTGTGTGATGTTTGGAAAGAGCCCTGTCTCTTCCACAGGCCAATCT	3000
Db	2941	GCCGGAAGCAACGTGTGTGTGATGTTTGGAAAGAGCCCTGTCTCTTCCACAGGCCAATCT	3000
QY	3001	CCATCTTACATTTGCTGCAACCAACATCTCTCAGATGAGTGTGATGATGAAGTGTGCG	3060
Db	3001	CCATCTTACATTTGCTGCAACCAACATCTCTCAGATGAGTGTGATGATGAAGTGTGCG	3060
QY	3061	GTCGAGGTGACAGGGCCAGATCCACACAGGACCTGGATCTTCAGTATGTGGAAAGACCCC	3120
Db	3061	GTCGAGGTGACAGGGCCAGATCCACACAGGACCTGGATCTTCAGTATGTGGAAAGACCCC	3120
QY	3121	ACCATGTGCGGATTTGACCCAGATGAGCATTTGTCAGTGAGAAACACACCCATGCGCGTA	3180
Db	3121	ACCATGTGCGGATTTGACCCAGATGAGCATTTGTCAGTGAGAAACACACCCATGCGCGTA	3180
QY	3181	TGGGGGAGCCCACTGGAACCTCATACAAACCCTGGATCTGTGCCAAGCATGAGAGGAG	3240
Db	3181	TGGGGGAGCCCACTGGAACCTCATACAAACCCTGGATCTGTGCCAAGCATGAGAGGAG	3240
QY	3241	GAGCAATCAATATCTGTGAGGTTCTGAACGCTTCTAGATGATGACTGTACGGCCGCGC	3300
Db	3241	GAGCAATCAATATCTGTGAGGTTCTGAACGCTTCTAGATGATGACTGTACGGCCGCGC	3300
QY	3301	CTCGCTCTGGGATCTTGACCAACATCAAGACTTGACCCAGAGGCCCGAGAGTTTGGCTTC	3360
Db	3301	CTCGCTCTGGGATCTTGACCAACATCAAGACTTGACCCAGAGGCCCGAGAGTTTGGCTTC	3360
QY	3361	ATCTCTGACAAAGTCCAGTCCCTGTGCTCATCTCTCAACAAAGACCACTTCACTACTATCC	3420
Db	3361	ATCTCTGACAAAGTCCAGTCCCTGTGCTCATCTCTCAACAAAGACCACTTCACTACTATCC	3420
QY	3421	AACCCGGGTTTGAAGCTTTGGTCCCTCCAGGAATCTGTGAGTCAAGCCCTGGACAGCC	3480
Db	3421	AACCCGGGTTTGAAGCTTTGGTCCCTCCAGGAATCTGTGAGTCAAGCCCTGGACAGCC	3480
QY	3481	ATCATCTTAAAGGCAAGACTGATATCCGCTGTGTGGCTGTGGGGCAACGTGAACTGAA	3540

Db 3481 ATCATCTTAAGGCGAAGAACCTGATCCCGCTGTGGCTGGGGGCAACGTGAAGCTGAAC 3540
Qy 3541 TACACTGTGTGTTGGGAGAGACCGTGAACCGTGAACCGTGTCAATGTCAAGCTGCTC 3600
Db 3541 TACACTGTGTGTTGGGAGAGACCGTGAACCGTGTCAATGTCAAGCTGCTC 3600
Qy 3601 TGGAGTGTCCCAACCTCATCGGAGGCAAAAGTGAAGCCGTGTGGTGGATGAG 3660
Db 3601 TGGAGTGTCCCAACCTCATCGGAGGCAAAAGTGAAGCCGTGTGGTGGATGAG 3660
Qy 3661 TACTTCCCGGGGATGTGTATCTTCCCGGACAGCCCGCTCAAGCTCCCGCATGTC 3720
Db 3661 TACTTCCCGGGGATGTGTATCTTCCCGGACAGCCCGCTCAAGCTCCCGCATGTC 3720
Qy 3721 AGCATCGAGTGTGGGCGGCTCTCATATTTCATCGTGGCCGTGTCAATGGCTAT 3780
Db 3721 AGCATCGAGTGTGGGCGGCTCTCATATTTCATCGTGGCCGTGTCAATGGCTAT 3780
Qy 3781 AAAGCGAAGTCTCCCGGAAGTGAACCTCAAGTGAAGGCTGTCAATGTGAAC 3840
Db 3781 AAAGCGAAGTCTCCCGGAAGTGAACCTCAAGTGAAGGCTGTCAATGTGAAC 3840
Qy 3841 CTGAGTGTCCGTGTGGCCCTGAGTGAAGAAAGCTTTGCGAGCTGAGACGAGATC 3900
Db 3841 CTGAGTGTCCGTGTGGCCCTGAGTGAAGAAAGCTTTGCGAGCTGAGACGAGATC 3900
Qy 3901 CATAGAGTGAACCTGATGATGAGAGCCGGATTCCTTCTGAGCTATGAACTTAC 3960
Db 3901 CATAGAGTGAACCTGATGATGAGAGCCGGATTCCTTCTGAGCTATGAACTTAC 3960
Qy 3961 ACCATGCGAGTGTGTTCCTCCAGAAATTGAAGACCAACCTTCTCCGGAGCTTGAAGTC 4020
Db 3961 ACCATGCGAGTGTGTTCCTCCAGAAATTGAAGACCAACCTTCTCCGGAGCTTGAAGTC 4020
Qy 4021 CCGGGCTACCGGAGAGCGTGTGAGAAAGGCTGAAAGCTTTCCGACCTCAATCAAC 4080
Db 4021 CCGGGCTACCGGAGAGCGTGTGAGAAAGGCTGAAAGCTTTCCGACCTCAATCAAC 4080
Qy 4081 AACAGAGTGTCTGCTGTCTCTTCATCCGACGCTTGAAGTCCAGAGTGTGATCTTCATG 4140
Db 4081 AACAGAGTGTCTGCTGTCTCTTCATCCGACGCTTGAAGTCCAGAGTGTGATCTTCATG 4140
Qy 4141 CGGAGCGTGGAGAGCGTGTGAGAAAGGCTGAAAGCTTTCCGACCTCAATCAAC 4200
Db 4141 CGGAGCGTGGAGAGCGTGTGAGAAAGGCTGAAAGCTTTCCGACCTCAATCAAC 4200
Qy 4201 GCCACTGATGTGTGAAGCAAGCTGTGCGGACCTCAATTGACAAAGACCTGAGAGCAAG 4260
Db 4201 GCCACTGATGTGTGAAGCAAGCTGTGCGGACCTCAATTGACAAAGACCTGAGAGCAAG 4260
Qy 4261 AACCAACCTGAGTGTGTGAGAGCAAGCTGTGCGGACCTCAATTGACAAAGACCTGAGAGCAAG 4320
Db 4261 AACCAACCTGAGTGTGTGAGAGCAAGCTGTGCGGACCTCAATTGACAAAGACCTGAGAGCAAG 4320
Qy 4321 TGGTTACTTTCCTCTCTCAAGTTCTCAAGAGTGTCTGGGAGGCCCCCTCTTCTC 4380
Db 4321 TGGTTACTTTCCTCTCTCAAGTTCTCAAGAGTGTCTGGGAGGCCCCCTCTTCTC 4380
Qy 4381 CTGTTCTGTGCAATCAAGAGCAAGCTGTGCGGACCTCAATTGACAAAGACCTGAGAGCAAG 4440
Db 4381 CTGTTCTGTGCAATCAAGAGCAAGCTGTGCGGACCTCAATTGACAAAGACCTGAGAGCAAG 4440
Qy 4441 GCCCGCTACTCTTGAAGAGAGCAAGCTCAATCCGACAGATTGAATGAACAAACCTTG 4500
Db 4441 GCCCGCTACTCTTGAAGAGAGCAAGCTCAATCCGACAGATTGAATGAACAAACCTTG 4500
Qy 4501 GTCTGAGCTGTGTCAAGCCAGCAATGTGCAAGCCGAGGTCTCCAGTAAAGATCTC 4560
Db 4501 GTCTGAGCTGTGTCAAGCCAGCAATGTGCAAGCCGAGGTCTCCAGTAAAGATCTC 4560
Qy 4561 AACTGTGACCAATCACTCAAGTCAAGAGAAATTGTGAATGCCATTTCAAGAAATGTG 4620

Db 4561 AACTGTGACCAATCACTCAAGTCAAGAGAAATTGTGAATGCCATTTCAAGAAATGTG 4620
Qy 4621 CTTTGTCTCCACCGGCGCAAAAGTGTGAATGTGATCTGAGATGGCGCAAGAAAGTGG 4680
Db 4621 CTTTGTCTCCACCGGCGCAAAAGTGTGAATGTGATCTGAGATGGCGCAAGAAAGTGG 4680
Qy 4681 GCAAGATGATCTTGAAGATGAAGCAATCAACCAAGATTGAATGATTTGAAGGCA 4740
Db 4681 GCAAGATGATCTTGAAGATGAAGCAATCAACCAAGATTGAATGATTTGAAGGCA 4740
Qy 4741 CTGAACACTGGCCCACTACAGAGTGCAGATGTGATCCGAGTGGCAATGATGTCAG 4800
Db 4741 CTGAACACTGGCCCACTACAGAGTGCAGATGTGATCCGAGTGGCAATGATGTCAG 4800
Qy 4801 CAGTGAAGCTCTATACAGCAAGTGAACCACTCAACCTCTCCAGAGCTCAAGATGAA 4860
Db 4801 CAGTGAAGCTCTATACAGCAAGTGAACCACTCAACCTCTCCAGAGCTCAAGATGAA 4860
Qy 4861 TATGAATAATGATCCGTAACAGGAGAGCCCGGACAGCTCCGTAAGAGACCTATG 4920
Db 4861 TATGAATAATGATCCGTAACAGGAGAGCCCGGACAGCTCCGTAAGAGACCTATG 4920
Qy 4921 ATCACTCTGAACCTGAGAGTGAAGTGAATGATGTGCACTTGAAGAACCAAGAGAC 4980
Db 4921 ATCACTCTGAACCTGAGAGTGAAGTGAATGATGTGCACTTGAAGAACCAAGAGAC 4980
Qy 4981 GGAAGCAAG 5040
Db 4981 GGAAGCAAG 5040
Qy 5041 CTCCTGGCACTAAGGCACTGTGAAGATTTGTGATGATCTTGTGAACCAATCTTC 5100
Db 5041 CTCCTGGCACTAAGGCACTGTGAAGATTTGTGATGATCTTGTGAACCAATCTTC 5100
Qy 5101 AGAAGGCAACCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 5160
Db 5101 AGAAGGCAACCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 5160
Qy 5161 GATGAG 5220
Db 5161 GATGAG 5220
Qy 5221 AATTGCTGCTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 5280
Db 5221 AATTGCTGCTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 5280
Qy 5281 ATCCATTAAGAAAGCAAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5340
Db 5281 ATCCATTAAGAAAGCAAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5340
Qy 5341 TCTTGTCCACGTCAG 5400
Db 5341 TCTTGTCCACGTCAG 5400
Qy 5401 GCCAAGAGATCCCGCACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5460
Db 5401 GCCAAGAGATCCCGCACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5460
Qy 5461 ATCCAGAGATCCAG 5520
Db 5461 ATCCAGAGATCCAG 5520
Qy 5521 ATGAATGATGATCAACATGATGATGATGATGATGATGATGATGATGATGATGATG 5580
Db 5521 ATGAATGATGATCAACATGATGATGATGATGATGATGATGATGATGATGATGATG 5580
Qy 5581 AGGAGAGATCTTGAAGATCTTGAAGATCTTGAAGATCTTGAAGATCTTGAAGATCTT 5640
Db 5581 AGGAGAGATCTTGAAGATCTTGAAGATCTTGAAGATCTTGAAGATCTTGAAGATCTT 5640
Qy 5641 TACAACTGAACAGTCAATACCTCATGAGCTTGAAGAGAGAGAGAGAGAGAGAGAG 5691
Db 5641 TACAACTGAACAGTCAATACCTCATGAGCTTGAAGAGAGAGAGAGAGAGAGAGAG 5691

RESULT 2
ABA00062
ID ABA00062 standard; cDNA, 6367 BP.
XX
XX ABA00062;
XX
XX 25-OCF-2002 (first entry)
XX
XX CADHP-9 coding sequence, Incyte ID No: 7156379CBL.
XX
XX Gene; human; cell adhesion protein; CADHP; AIDS; Alzheimer's disease;
XX acquired immunodeficiency syndrome; thymic dysplasia; epilepsy;
XX renal tubular acidosis; congenital glaucoma; cancer; atherosclerosis;
XX Parkinson's disease; ss.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX CDS 566..6250
XX /tag= a
XX /product= "CADHP-9"
XX
XX W0200259312-A2.
XX
XX 01-AUG-2002.
XX
XX 18-DEC-2001; 2001MO-US049206.
XX
XX 18-DEC-2000; 2000US-0256542P.
XX 22-DEC-2000; 2000US-0259604P.
XX 05-JAN-2001; 2001US-0260101P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Duggan BM, Xu Y, Lee S, Lu DM, Warren BA, Yue H;
XX Gietzen KJ, Honchell EA, Burford N, Baughn MR, Tang TY, Hillman JL;
XX Gandhi AR, Kallik DA, Bandman O, Grail RC, Walia NK, Lu Y;
XX Ramkumar J, Yeo MG, Lal PG;
XX WPI; 2002-590826/63.
XX P-PSDB; AAG79420.
XX
XX New human cell adhesion proteins (CADHP) useful for treating, diagnosing
XX and preventing diseases or conditions associated with the aberrant CADHP
XX expression e.g. cancer, acquired immunodeficiency syndrome, Alzheimer's
XX disease and epilepsy.
XX
XX Claim 5; Page 146-48; 149pp; English.
XX
XX The sequences given in ABA00054-63 encode novel human cell adhesion
XX proteins (CADHP). The CADHP polypeptides and polynucleotides are useful
XX in treating, diagnosing and preventing diseases or conditions associated
XX with the decreased expression or overexpression of CADHP, e.g. immune
XX system (acquired immunodeficiency syndrome, thymic dysplasia),
XX neurological (Alzheimer's disease, Parkinson's disease, epilepsy),
XX developmental (renal tubular acidosis, congenital glaucoma) and cell
XX proliferative (cancer, atherosclerosis) disorders. They are also useful
XX in assessing the effects of exogenous compounds on the expression of
XX nucleic acid and amino acid sequences of CADHP. The CADHP or its
XX fragments are useful in screening compounds for effectiveness as agonist
XX or antagonist of the polypeptides, or in altering the expression of the
XX target polynucleotide and compounds that specifically bind to or modulate
XX the activity of the polypeptide. The protein encoded by this cDNA
XX sequence shows homology to mouse plexin-2
XX
XX Sequence 6367 BP; 1423 A; 1885 C; 1799 G; 1259 T; 0 U; 1 Other;
XX
XX Query Match 99.8%; Score 5680.6; DB 6; Length 6367;
XX Best Local Similarity 99.9%; Pired. No. 0;
XX Matches 5683; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGAAGCCATGCGCTGGAACTGGAACCTGCTCTCTCCACCTCTCATGTGGGCAATG 60
Db 566 ATGAAGCCATGCGCTGGAACTGGAACCTGCTCTCTCCACCTCTCATGTGGGCAATG 625
QY 61 GGCCTCTCCATCTTGTCTCACCCTGGAGACCCGCTGTCCAGAGAGAGAGAGAGAGAG 120
Db 626 GGCCTCTCCATCTTGTCTCACCCTGGAGACCCGCTGTCCAGAGAGAGAGAGAGAGAG 685
QY 121 GTCAATTCGAG 180
Db 686 GTCAATTCGAG 745
QY 181 GACACATTTTACTTGGGGGCGCTCATCGATTTTACAGCTCTCCAGGACCTGAGAGTTC 240
Db 746 GACACATTTTACTTGGGGGCGCTCATCGATTTTACAGCTCTCCAGGACCTGAGAGTTC 805
QY 241 TTGGTGAAGCATGAG 300
Db 806 TTGGTGAAGCATGAG 865
QY 301 GTCCAGACCTGAG 360
Db 866 GTCCAGACCTGAG 925
QY 361 GACTACAG 420
Db 926 GACTACAG 985
QY 421 CTGAGGCTGAG 480
Db 986 CTGAGGCTGAG 1045
QY 481 TCAGGTGTCAAG 540
Db 1046 TCAGGTGTCAAG 1105
QY 541 GACAGAGCTGTCATATGTCACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db 1106 GACAGAGCTGTCATATGTCACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1165
QY 601 AGCCGAAAGTACCAAG 660
Db 1166 AGCCGAAAGTACCAAG 1225
QY 661 GAGTTCGAGCTGAG 720
Db 1226 GAGTTCGAGCTGAG 1285
QY 721 GATATCTACTATGCTATGCTTTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db 1286 GATATCTACTATGCTATGCTTTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1345
QY 781 CCTGAGATGAGTGTCTTCCACAGAGCTCCACACCAAGAGAGAGAGAGAGAGAGAGAG 840
Db 1346 CCTGAGATGAGTGTCTTCCACAGAGCTCCACACCAAGAGAGAGAGAGAGAGAGAGAG 1405
QY 841 GTGAGGCTTTGAG 900
Db 1406 GTGAGGCTTTGAG 1465
QY 901 GAGGCGAG 960
Db 1466 GAGGCGAG 1525
QY 961 GTGCTTGGAG 1020
Db 1526 GTGCTTGGAG 1585
QY 1021 AAGGCGAG 1080
Db 1586 AAGGCGAG 1645
QY 1081 AAGCAGATTAATGACCGCATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140

1646 AAGCAGTAAATGACCGATTAAGAGCGGCTGAGTCTTTGATCCGAGGCGAGGGACG 1705
1141 CTGACCTGGGCTGCTCAAGGTGAAGAGACATCCCTGACAGAGTGGCTTTAAACAT 1200
1706 CTGACCTGGGCTGCTCAAGGTGAAGAGACATCCCTGACAGAGTGGCTTTAAACAT 1765
1201 GACAGTAACTCTGTGGCTGAGCATGATGCTCCCTGGAGATGCCAGATGGTGGT 1260
1766 GACAGTAACTCTGTGGCTGAGCATGATGCTCCCTGGAGATGCCAGATGGTGGT 1825
1261 GGAATTCCTGCTTCAAGGAGACAGGAGCCGATGACGTCTGATCGCATATGCTAC 1320
1826 GGAATTCCTGCTTCAAGGAGACAGGAGCCGATGACGTCTGATCGCATATGCTAC 1885
1321 AAGAAACCACTCTGGGCTTTGTGGGACCAAAAGTGGCAAGCTGAAGAAAGATCCGGGTG 1380
1886 AAGAAACCACTCTGGGCTTTGTGGGACCAAAAGTGGCAAGCTGAAGAAAGATCCGGGTG 1945
1381 GATGAGCCAGGAGCAAGCCCTTCAGTATGAGCGGTGAGGCTGGTGGATCCCGGCTCA 1440
1946 GATGAGCCAGGAGCAAGCCCTTCAGTATGAGCGGTGAGGCTGGTGGATCCCGGCTCA 2005
1441 GTCCCTCCGAGATGAGCTTCTCCAGAGACACAGAGCACTTACATCATGTCAGAGAG 1500
2006 GTCCCTCCGAGATGAGCTTCTCCAGAGACACAGAGCACTTACATCATGTCAGAGAG 2065
1501 CAGCTACACAGATGCTCTGTGAGTCTGTGTGATGATGAGAGCTGGCGGAGTGGCTT 1560
2066 CAGCTACACAGATGCTCTGTGAGTCTGTGTGATGATGAGAGCTGGCGGAGTGGCTT 2125
1561 GAGTCAAGGAGCAAGCCCTTCAGTATGAGCGGTGAGGCTGGTGGATCCCGGCTCA 1620
2126 GAGTCAAGGAGCAAGCCCTTCAGTATGAGCGGTGAGGCTGGTGGATCCCGGCTCA 2185
1621 CGGTGTGAGCGGTGTCAGAGAGCCCGGAGGTTGGCTCGAGATGAGAGTGTGTCCG 1680
2186 CGGTGTGAGCGGTGTCAGAGAGCCCGGAGGTTGGCTCGAGATGAGAGTGTGTCCG 2245
1681 CTGACGCTGATCCCAACAATATCTCCGCTCTCAGTACAGAGTGTGTGTCTGAG 1740
2246 CTGACGCTGATCCCAACAATATCTCCGCTCTCAGTACAGAGTGTGTGTCTGAG 2305
1741 ACGTACATATGCTCCGAGAGTGTCACTGAGTCACTGACCTTTGAGAGCTGTGAG 1800
2306 ACGTACATATGCTCCGAGAGTGTCACTGAGTCACTGACCTTTGAGAGCTGTGAG 2365
1801 ATGATGGGCTGTGCTGGGCAATCAGATCAGTGTACTCCCTGACAGCCAGAGAGT 1860
2366 ATGATGGGCTGTGCTGGGCAATCAGATCAGTGTACTCCCTGACAGCCAGAGAGT 2425
1861 CCGCGGATCATCAGAGAAATGGGAGCAACATGTGCTACAGTTCAAGCTCAATTAAG 1920
2426 CCGCGGATCATCAGAGAAATGGGAGCAACATGTGCTACAGTTCAAGCTCAATTAAG 2485
1921 GAGACCGGATGACCTTGGCCAGACAGAGCTTTGTCTTACAAATTGACAGCTCAAT 1980
2486 GAGACCGGATGACCTTGGCCAGACAGAGCTTTGTCTTACAAATTGACAGCTCAAT 2545
1981 TGGTCCCTGTCTGTGGTGAAGTGCATACCGGCTGACATGTGTAAATACCGGCTGTC 2040
2546 TGGTCCCTGTCTGTGGTGAAGTGCATACCGGCTGACATGTGTAAATACCGGCTGTC 2605
2041 TGCACCATGACCCCAAGACCTGTCTTCCAGAGAGCCGAGTGAAGCTGCCAGAGAC 2100
2606 TGCACCATGACCCCAAGACCTGTCTTCCAGAGAGCCGAGTGAAGCTGCCAGAGAC 2665
2101 TGGCCCGAGCTGTGGGAGTGAAGATCTGTGTGCGCGGAGAGGATCAAGCTTAC 2160
2666 TGGCCCGAGCTGTGGGAGTGAAGATCTGTGTGCGCGGAGAGGATCAAGCTTAC 2725
2161 ACGCTGAAGGCCAAGACCTTCCCGAGCCCAAGTGTGGGAGCGGTGCTACGAATCAT 2220

2726 ACGCTGAAGGCCAAGACCTCCCGAGGCCAGGTGTGGGCAAGCGTGTACGAATGATC 2785
2221 CTCAATCATGAGGAGCAGAGAGTGGTCCCGCTGCGCTTAAACAGCTCCAGCTA 2280
2786 CTCAATCATGAGGAGCAGAGAGTGGTCCCGCTGCGCTTAAACAGCTCCAGCTA 2845
2281 CAGTGCAGAAACCTCTTATCTATGAGAGGATGAGATCAACACTGCCGTGGAG 2340
2846 CAGTGCAGAAACCTCTTATCTATGAGAGGATGAGATCAACACTGCCGTGGAG 2905
2341 TTGACAGTGTGTGGAATGGGCACTTCAACATTTGACCAACCCAGCTCAGATTAAGTTAC 2400
2906 TTGACAGTGTGTGGAATGGGCACTTCAACATTTGACCAACCCAGCTCAGATTAAGTTAC 2965
2401 CTCTAAGTGTGAGCCATGCGTGAAGCTGCGGGCTGTGCTCAAGCTGACCCAGAC 2460
2966 CTCTAAGTGTGAGCCATGCGTGAAGCTGCGGGCTGTGCTCAAGCTGACCCAGAC 3025
2461 TTGGCATGTGGTGGTCCAGAGGCCCAAGCCAGTGCACCTGGCGCAGACCTGGCC 2520
3026 TTGGCATGTGGTGGTCCAGAGGCCCAAGCCAGTGCACCTGGCGCAGACCTGGCC 3085
2521 CAGAGAGCCAGTGTGAGGCTGTGTGTGTCGCAAAAGCAAGTGCACAAACCCCGCATC 2580
3086 CAGAGAGCCAGTGTGAGGCTGTGTGTGTCGCAAAAGCAAGTGCACAAACCCCGCATC 3145
2581 ACAGAGTAAATCCCGGTGACAGGCCCCCGGAGAGGGGACCAAGAGTCACTATCCAGAG 2640
3146 ACAGAGTAAATCCCGGTGACAGGCCCCCGGAGAGGGGACCAAGAGTCACTATCCAGAG 3205
2641 GAGAACCTGGGCTGGAATTTGCGCATGCTGCTCCATGTCAAGTGTGTGGCTGGAG 2700
3206 GAGAACCTGGGCTGGAATTTGCGCATGCTGCTCCATGTCAAGTGTGTGGCTGGAG 3265
2701 TGGAGCCCTTGAAGTGTGATCAATCCCTGAGAGACAGATGATGTGAGTGGGAG 2760
3266 TGGAGCCCTTGAAGTGTGATCAATCCCTGAGAGACAGATGATGTGAGTGGGAG 3325
2761 GCCAAGCCAGCCAGCATCAGAGCTTGTGAGAGTGTGGTGGTGTGTGGCTGAA 2820
3326 GCCAAGCCAGCCAGCATCAGAGCTTGTGAGAGTGTGGTGGTGTGTGGCTGAA 3385
2821 TTGATGGCCCGGTCCCAAGCTCTATTACTTACATGACAGTCTCTCAGATCTGAG 2880
3386 TTGATGGCCCGGTCCCAAGCTCTATTACTTACATGACAGTCTCTCAGATCTGAG 3445
2881 CCGAGCCGAGGCGCATGTCCGAGGAGCCAGATGACCATCAAGCAACCTGAAT 2940
3446 CCGAGCCGAGGCGCATGTCCGAGGAGCCAGATGACCATCAAGCAACCTGAAT 3505
2941 GCCGAGAGCAAGTGTGTGATGTTTGAAGAGAGCCCTGTCTTCAAGAGGAGTCT 3000
3506 GCCGAGAGCAAGTGTGTGATGTTTGAAGAGAGCCCTGTCTTCAAGAGGAGTCT 3565
3001 CCAATCTCAATTTGTCTGACACACACATCTCAATGAGTGTGATGAATGAAGTGTG 3060
3566 CCAATCTCAATTTGTCTGACACACACATCTCAATGAGTGTGATGAATGAAGTGTG 3625
3061 GTGAGGTGAGCAGGAGCCAAATCCACAGAGCTGTCTTCAATGATGAGAGAGCC 3120
3626 GTGAGGTGAGCAGGAGCCAAATCCACAGAGCTGTCTTCAATGATGAGAGAGCC 3685
3121 ACCATGTGTGGATTTGAGCCAGATGAGATTTGTCAGTGGAACCAACCCATGCCGTA 3180
3686 ACCATGTGTGGATTTGAGCCAGATGAGATTTGTCAGTGGAACCAACCCATGCCGTA 3745
3181 TGGGAGAGCCAGCTGAGCTCATACAGAAACCCCGAGTCCGCAAGATGAGAGAG 3240
3746 TGGGAGAGCCAGCTGAGCTCATACAGAAACCCCGAGTCCGCAAGATGAGAGAG 3805
3241 GAGACATCAATATCTGTGAGGCTTGAAGAGCTGATGAGATCACTGTGAGGCGCCGCC 3300
3806 GAGACATCAATATCTGTGAGGCTTGAAGAGCTGATGAGATCACTGTGAGGCGCCGCC 3865

QY	3301	CTCGCTCTGAGGCTCTGACACCAAGTGAAGCTTGAACGAGAGGCGCGAGGATGTTTGGCTTC	3360
Db	3866	CTTGCTCTTGAGGCTCTTGACCAACCAATCAACCTTGACCGAGAGGCGCGAGGATGTTTGGCTTC	3925
QY	3361	ATCTTGAGCAACGTCACAGTCCCTCTGCTCATCTCTCAACAGACCAACTTCACTACTATCCC	3420
Db	3926	ATCTTGAGCAACGTCACAGTCCCTCTGCTCATCTCTCAACAGACCAACTTCACTACTATCCC	3985
QY	3421	AACCCGGTGTTTGAGGCGCTTTGGTTCCTCCAGAGATCCTTGAGAGTCAAGGCTGGCAAGGCC	3480
Db	3986	AACCCGGTGTTTGAGGCGCTTTGGTTCCTCCAGAGATCCTTGAGAGTCAAGGCTGGCAAGGCC	4045
QY	3481	ATCATCTTAAAGGCGCAAGAACCTGTATCCCGCTGTGGCTTGGGGCAACGCTGAAGCTGAAC	3540
Db	4046	ATCATCTTAAAGGCGCAAGAACCTGTATCCCGCTGTGGCTTGGGGCAACGCTGAAGCTGAAC	4105
QY	3541	TACACTGTACTGTGTTGGGGAGAAAGCGTGACACGCTGACCGTGTCAATGTTCACAGCTGTCTC	3600
Db	4106	TACACTGTACTGTGTTGGGGAGAAAGCGTGACACGCTGACCGTGTCAATGTTCACAGCTGTCTC	4155
QY	3601	TGGCACTCCCCCAACCTCATGTGGCAGGCGACAAAGTGAATGGCCCGTGTGCTGTGGCATGAG	3660
Db	4166	TGGCACTCCCCCAACCTCATGTGGCAGGCGACAAAGTGAATGGCCCGTGTGCTGTGGCATGAG	4225
QY	3661	TACTCCCGGGGGATGGTGTACATTGGCCCCGGGACAGCCCGTCTAGCGCTGGCCGSCATCTGC	3720
Db	4226	TACTCCCGGGGGATGGTGTACATTGGCCCCGGGACAGCCCGTCTAGCGCTGGCCGSCATCTGC	4285
QY	3721	AGCATGCAGTGACCTGGCGGCGCTCTCTCATCTTTTCATGTGGCCGCTGTCAATGCTCTAT	3780
Db	4286	AGCATGCGGATGGCTGGCGGCGCTCTCTCTCATCTTTTCATGTGGCCGCTGTCAATGCTCTAT	4345
QY	3781	AAACGCAATCTCCCGGAGAAAGTGAACCTCATCGCTGAAGGCTGTGAGATGAGATGAGACAAC	3840
Db	4346	AAACGCAATCTCCCGGAGAAAGTGAACCTCATCGCTGAAGGCTGTGAGATGAGATGAGACAAC	4405
QY	3841	CTGGAATCCCTCGTGTGGGCGCTGTGAGTGAAGGAACCTTTTCCGAGCTGTGAACGCAATTC	3900
Db	4406	CTGGAATCCCTCGTGTGGGCGCTGTGAGTGAAGGAACCTTTTCCGAGCTGTGAACGCAATTC	4465
QY	3901	CATGAGCTGACACAGTGAACCTGGATGTGAACCGGAGATTCGTTCTCTGGACATATGAATTCAC	3960
Db	4466	CATGAGCTGACACAGTGAACCTGGATGTGAACCGGAGATTCGTTCTCTGGACATATGAATTCAC	4525
QY	3961	ACCATGCGGAGTGTCTGTTCCCAAGAAATTGAAGACCACTCTGTCTCTCCGGAACCTTTAGAGTTC	4020
Db	4526	ACCATGCGGAGTGTCTGTTCCCAAGAAATTGAAGACCACTCTGTCTCTCCGGAACCTTTAGAGTTC	4585
QY	4021	CCGGGAGTCAACGGAGAGAGGTGTGGAAGAAAGCTGTGAAGCTTTTCCGCAAGCTCATCAAC	4080
Db	4586	CCGGGAGTCAACGGAGAGAGGTGTGGAAGAAAGCTGTGAAGCTTTTCCGCAAGCTCATCAAC	4645
QY	4081	AACAGAGTGTCTGTGTCTCTTTCATCCGACGCTTGAATGCCAGGTGATGCTTCTTCATG	4140
Db	4646	AACAGAGTGTCTGTGTCTCTTTCATCCGACGCTTGAATGCCAGGTGATGCTTCTTCATG	4705
QY	4141	CGGAGACCGTGGCAACGTGGAGCTCATCTCATATGACCGTGTGTGAGAGCAAGCTTGAGATAC	4200
Db	4706	CGGAGACCGTGGCAACGTGGAGCTCATCTCATATGACCGTGTGTGAGAGCAACGTTGAGATAC	4755
QY	4201	GCCACTGATGTGTGTAAGCACTGCTGGCGGACCTCATTTGAACAAGAACTTGAGAGCAAG	4260
Db	4766	GCCACTGATGTGTGTAAGCACTGCTGGCGGACCTCATTTGAACAAGAACTTGAGAGCAAG	4825
QY	4261	AACCACTCTAAGCTGTGTCTCAGAGAGCATGATCACTGGCTAGAGAGATGTCTGCACAT	4320
Db	4826	AACCACTCTAAGCTGTGTCTCAGAGAGCATGATCACTGGCTAGAGAGATGTCTGCACAT	4885
QY	4321	TGGTTTAACTTTCTCTCTCTTCAAAAGTTCTCTCAAGAGATGTCTGGGAGAGGCCCTCTTCTCC	4380
Db	4886	TGGTTTAACTTTCTCTCTCTTCAAAAGTTCTCTCAAGAGATGTCTGGGAGAGGCCCTCTTCTCC	4945

QY	4381	CTGTTCTGGCCATTAACACACAAATGGAAGAAAGGCCCCATTGAGCCATCAACGGGCGAG	44440
Db	4346	CTGTTCTGGCCATTAACACACAAATGGAAGAAAGGCCCCATTGAGCCATTAACGGGCGAG	5005
QY	4441	GCCCGCTACTCCTTGAGCGAAGCAAGCTCATCCGCGAGAGATTGACTACAAACCTCG	4500
Db	5006	GCCCGCTACTCCTTGAGCGAAGCAAGCTCATCCGCGAGAGATTGACTACAAACCTCG	5065
QY	4501	GTCTCAGTGTGTGAGCGCCAGACAAATGAGCCAGCCCGAGGTCCAGTAAAGATCTCTC	4560
Db	5066	GTCTCAGTGTGTGAGCGCCAGACAAATGAGCCAGCCCGAGGTCCAGTAAAGATCTCTC	5125
QY	4561	AACGTGACCAATCACTCAGGTCAGAGAGAAATTTGTGATGCAATCTTCAAGATGTG	4620
Db	5126	AACGTGACCAATCACTCAGGTCAGAGAGAAATTTGTGATGCAATCTTCAAGATGTG	5185
QY	4621	CCCTGTCCCAACGGGCCCCAAGCTGCGAGATATCTGAGAGGGGCAACAAGAAATGTGG	4680
Db	5186	CCCTGTCCCAACGGGCCCCAAGCTGCGAGATATCTGAGAGGGGCAACAAGAAATGTGG	5245
QY	4681	GCAAGGATATCTTGACAGATGAAGACATCAACAAGATTGAGATGATTGGAAGCGA	4740
Db	5246	GCAAGGATATCTTGACAGATGAAGACATCAACAAGATTGAGATGATTGGAAGCGA	5305
QY	4741	CTGAACACACTGGGCCCATACAGAGTCCAGATCGTTCGTGATGAGCATTAAGTGCACAG	4800
Db	5306	CTGAACACACTGGGCCCATACAGAGTCCAGATCGTTCGTGATGAGCATTAAGTGCACAG	5365
QY	4801	CAGGTGACAGCTTATACGCACTGTAACAATCCACCTGCTCCAGACCTCAGACAGTAA	4860
Db	5366	CAGGTGACAGCTTATACGCACTGTAACAATCCACCTGCTCCAGACCTCAGACAGTAA	5425
QY	4861	TATGAAAACATGATCCGTTACACGAGCGAGCCCGACAGGCTCCGTCACGACACCTATG	4920
Db	5426	TATGAAAACATGATCCGTTACACGAGCGAGCCCGACAGGCTCCGTCACGACACCTATG	5485
QY	4921	ATCACTCTCTGACTGAGAGATGTGAGTCAAGATGTGGACCTTATGTGAAGAACCAAGAGAC	4980
Db	5486	ATCACTCTCTGACTGAGAGATGTGAGTCAAGATGTGGACCTTATGTGAAGAACCAAGAGAC	5545
QY	4981	GGAGACCAAGAGAGGGGAGCCGAGGAGAGAAATGATGTTGAAATCACTGACCCGCA	5040
Db	5546	GGAGACCAAGAGAGGGGAGCCGAGGAGAGAAATGATGTTGAAATCACTGACCCGCA	5605
QY	5041	CTCTGAGCCACTAAGGGGCAACTGACAGAAATTTGTGATGACCTCTTGAAGCAATCTTC	5100
Db	5606	CTCTGAGCCACTAAGGGGCAACTGACAGAAATTTGTGATGACCTCTTGAAGCAATCTTC	5665
QY	5101	AGCAGGGCACACCGTGGCTGACCCCTGCCCTGACATCAAGTATGATGATCTTCCTG	5160
Db	5666	AGCAGGGCACACCGTGGCTGACCCCTGCCCTGACATCAAGTATGATGATCTTCCTG	5725
QY	5161	GATGAGCAGGCTGATTAACATGGAATTCATGACCCGCAAGTCCGCAATACCTGGAAGAGC	5220
Db	5726	GATGAGCAGGCTGATTAACATGGAATTCATGACCCGCAAGTCCGCAATACCTGGAAGAGC	5785
QY	5221	AATTGCGTCCCTGAGGTTTGTGGATCAATGATCAAGAACCGAGATTGTGTGTGAC	5280
Db	5786	AATTGCGTCCCTGAGGTTTGTGGATCAATGATCAAGAACCGAGATTGTGTGTGAC	5845
QY	5281	ATCATTAAGAACAGATCAAGAGAGCTGCTCTCTGTGTGTGCTCAAGCTTCAATGAGAC	5340
Db	5846	ATCATTAAGAACAGATCAAGAGAGCTGCTCTCTGTGTGTGCTCAAGCTTCAATGAGAC	5905
QY	5341	TCTTGCTCCACTGAGACACCGGCTGGGCAAGGATCGCCCTCCAAACAAGCTGCTGAT	5400
Db	5906	TCTTGCTCCACTGAGACACCGGCTGGGCAAGGATCGCCCTCCAAACAAGCTGCTGAT	5965
QY	5401	GCCAAAGCAATCCCAAGCTTAACAATAATGGGTGGAAGGATTAATCAAGACATAGAGAG	5460
Db	5966	GCCAAAGCAATCCCAAGCTTAACAATAATGGGTGGAAGGATTAATCAAGACATAGAGAG	6025
QY	5461	ATGCCAGCCATCAGGACCAAGACATTAAGCAATACCTGAGTGAAGATCCCGGATGAC	5520

QY	CTGCAAGCGCTGACCTGCTGTCACCAAGCGGGGGCGCTGTGACAGACCCCTTGAGTCCAT	967
Db	1633 CTGCAAGCGCTGTTACCTGGCCAGAGCTGGGACATCGTCGGCCAGAGCTTTCATATCAC	1692
QY	988 CCAAGATGATGACCTGCTCTTCAACGCTTCTTCACAGGCGCAGAACGCGAAATTAATCC	1047
Db	1693 AGCCAGAGAGAGATGATCTCTTGGCCATCTTCTCCAAAGGGGAGAGAGATACACACCCG	1752
QY	1048 CTGATAGAGCGGCGCCCTGTGACATCTTTCATCTTGAAGAGATTAATATGACCGATTAAGAG	1107
Db	1753 CCGAGATGACTCTCCCTGTGTCTTCCATTCGGGGCATCAACTTGGAGATCAAGAG	1812
QY	1108 CGGCTGACGCTTGTTCACCGGGGGCGAGGGGACGTGACCTTGGCTGTGGCTCAAGGTGAG	1167
Db	1613 CGCTGTGAGTCTGTGACCAAGGGCGAGGGGCACTGGAGCTCAACTGSGCTGTGGGGAGG	1872
QY	1168 GACATCCCCCTGCAGCAGTGGCTCTTAAACAATTGACATTACTTGTGGCTGTGACATG	1227
Db	1873 GACGTCCAGTGCACCAAGGCGCTGTGCCCATCATATTAATCTTGTGACCTGACATC	1932
QY	1228 AATGCTCCCCCTGGAGGTGGCCGACATGTGTGGTGAATTCGCTCTTCAACGAGACAGG	1287
Db	1933 AACCAAGCCCTGGAGAGCTCAACTCGATGGAGGCGCTGACCTGTACACACACAGCAGG	1992
QY	1288 GACCCGATGACGTCTGTCAATGCGATATGTCTTCAAGAACACATCTCTGGCTTTGTGGGC	1347
Db	1993 GACCGGATGACCTCTGTGGCTCTCCACCTTTCACACGGCTACACGGCTGTATTTTGTGGGG	2052
QY	1348 ACCAAAAGTGGCAGCTGAGAGAGATCCGGGTGATGACCCAGGGGGCAACGCTCTCGAG	1407
Db	2053 ACTTAAAGTGGCAAGCTGAAAAAGATTGGGGCCGACGATCCCCCATGATGTGGGGTCCAG	2112
QY	1408 TATGAGACGCTGAGGTGTGT---GGAACCCGGGCCATGCTCCGGGATATGTGCTCTTCC	1454
Db	2113 TACGAAATGTCTCTGTGTCTCAAGGACGGAAAGCCCATCTCCGGGACATGGCTTTCTCC	2172
QY	1465 AAGGACCAAGAACACTTACATCATGTCAAGAGGCGAGCTCACCGAGTCCCTGTGGAG	1524
Db	2173 ATTGATCAACGCTACCTGTACCTCACTGCTGTAAGAGACAGGTCAACAGGATCCCCCTGTAG	2232
QY	1525 TCCCTGTGTCAATATCAGAGCTGCGGCGAGTGCCTTGCTCAAGGCCAACCCCATGTGGC	1584
Db	2233 TCATGTAGAGAGATATCGACTTGTGGGAGTGGCTGAGCTCTGGGGACCCCTCATCTGTGGC	2292
QY	1585 TGTGTGTGTGTGACAAACGTCGCAACCGGCAAGGAGGGGTGTAAGCGTCCAAAGAGGCC	1644
Db	2293 TGTGTGTCCCTGCAACAATGTGTCTCCCGCAGGGACAAATGCCAAAGCTGGGAACTT	2352
QY	1645 CGCAGTTGTGCTCGAGAGATGAAGAGATGTGTCCGCTGACGCTCATCCCAACATATC	1704
Db	2353 AATCGATTGTGCTCCGAGCATCAAGCCAGTGTGAGCTTTCAGTGTATCCACGACGATC	2412
QY	1705 TCGCTCTCTCAGTACACAGCTGTGTGTCTCTGAGACGTACAAATGTCCCGAGCTGAC	1764
Db	2413 TCAGTATCTGAGACAGCCGGTGTGTTAGCCTGGTATGATGATGCTCTGATATCATCT	2472
QY	1765 GCTTGGGTCAACTGCAACTTTTGAAGACCTGTCAAGATGAGATGGGTGCTGTGGGCAT	1824
Db	2473 GCGGTGATGCTCTGTGCTTTTGGAACTGACAGAGGTGAGGGGCGAGGTGCCGAGAC	2532
QY	1825 CAGATCCAGTGTACTCCCTTGCAACGCCAAGAGGTGCCCGGATCATCAAGAGATGGG	1884
Db	2533 CAGGTCACTGCACTCTCACTCGGGCCCAAGATGT---CCCTGTCACTCCCGCTGATCA	2589
QY	1885 GACCAACATNTGCTACAGCTTACAGTCAAAATCAAGAGACCGGCAATGACTTCCGACGC	1944
Db	2590 GACTGTGTTGGCTGGAGCTCACGCTGAGGTCCAAAGGACAGGAAAGATTTTGTGACG	2649
QY	1945 ACCAGCTTGTCTTTCACATTTGAGGCGTCCACAAATTCGTGCTGTCTCTGTGGAGAT	2004
Db	2650 ACCGAGTTCAAGTTTAAACATGCAAGTCCCAACCACTGTGCTGTCTGTGCAACGC	2709

[illegible]

Db 3790 GTGATAGCAACCTGAGTTGAGTACATAGATGACCTCGGGTCCAGGCCATCGAGCCA 3849
Qy 3142 GAATGAGATGTTGAGTGAAGCAACACCCATGGCGGTATGGGGGAGCCCACTGAGCTC 3201
Db 3850 GAGTGGAGCATTTGCGAGTGGCCACACCCCTGACATCAAGGCTTCAACTGGATGTC 3909
Qy 3202 ATACGAACCCCCAGATCCGTGCGAAGCATGAGAGGAGCAATCAATATCTGTGAG 3261
Db 3910 ATTCGAGAGCCAGATCCGATCGAATTCATGAGCAAGAAATCTGTCAAATGTGTAA 3969
Qy 3262 GTTTCGAGCTACTGAGATGATACCTGTGAGGCGCCGCTCGCTGAGGTCCTGACAC 3321
Db 3970 GTTGTGAACACACACCCCTGACCTGCTGAGCACTCTGTGACAGGATACCGGCTT 4029
Qy 3322 CAGTGAAGCTGACGAGAGGCGGAGAGTGGCTTCACTCTGACAAAGTTCGCTC 3381
Db 4030 GGCTGGAACATGTGGAAGGCCCAAGTGAATTTGTGTTAACTATGTCATATCC 4089
Qy 3382 CTGCTCATCTCAACAGACCACTTCACTATCTATCCAAACCGGATTTGAGGCTTT 3441
Db 4090 TTGCTAATTACAGACCACTGATTAATCTATACCCCAACCCGACCTTGAATGCTT 4149
Qy 3442 GGTTCCTCAGGAATCTCTGAGCTCAGCTGCGACCGCCATCTCTAAAGGCAAGAC 3501
Db 4150 AACCTCTAGAGATCTTGGATCAAAAGCCAGATGCGCCATCTTGAAGGCAAAAC 4209
Qy 3502 CTGATCCCGCTGTGGCTGGGGGCAAGTGAAGCTGACATCACTGTGCTGTGGGAG 3561
Db 4210 CTCTGCTCTGCTCTGAGAG--GACCACCACTCACTGCTCATCGAGAG 4266
Qy 3562 AAGCGGTGACCGGTGACCGGTGACGATGTCAGTGTGCTGAGTCCGACCCCAACCTCAG 3621
Db 4267 ACCCTTGTGCTGTACCGGTATCTGAGACCCAGCTTCTTGTGAGGCTCCCACTCAGC 4326
Qy 3622 GGCAGGACCAAGTATGAGCCCGTGTGAGGAGTGAATCTCCCGGAGATGTGTAC 3681
Db 4327 GGGCAGACCAAGTATGATGTCAGTGGGCGGAGTGTGTTCTCGCTGCTGAGT 4386
Qy 3682 ATTGGCCCGAGACCGCTCAGCTGCGCCGCTGACGATCGACATCGAGTGTGAGC 3741
Db 4387 GTCATCTCAGACAGTGTGCTGACCTGTGACGCTTGTGACATCGGCGGCGGCGAC 4446
Qy 3742 CTCTCATCATTTTATGCTGAGCGGTGCTCATTTGCTTAAAGCAATCCGCGAAGT 3801
Db 4447 CTCTCTCATCATGATGATCATGATCTCTCATTTGCTTAAAGGCAAGTCTGAGAAAT 4506
Qy 3802 GACCTCAAGCTGAAGCGCTGACATGAGTGAAGCAACTGAGTCCGCTGAGCCCTG 3861
Db 4507 GACCTCACTCTCAAGCGGCTCAAAATGAGATGAGCAATCTGAGTCCGCTGAGCCCTG 4566
Qy 3862 GAGTGCAGAGAGCTTTGCGGAGCTGAGACGAGCATCATGAGCTGACATGACCTG 3921
Db 4567 GAGTGCAGAGAGCTTTGCTGAGCTCAGCGAGTATCAATGATTTGACCACTGACCTG 4626
Qy 3922 GATGAGCCGGATTCCTTCTGAGTATGAACTTACACATGCGGCTGTTCCTCA 3981
Db 4627 GACCGCTCAGGATCTCTTACCTGAGCTATGTAAGCTATGAGAGTCCGCTGCCG 4686
Qy 3982 GGAATTGAAGACACCGCTGCTCGGAGCTTGAAGTCCCGGCTACCGGAGAGGCT 4041
Db 4687 GGCATGAGAGACACCGGCTGCTGAGGAGCTGAGAGTCAAGAGACGGGACGAGCAC 4746
Qy 4042 GTGAGAGAAAGGCTTGAAGCTTTGCGCAGCTCATCAACAAAGAGTTCCTGCTGTC 4101
Db 4747 GTGAGAGAAAGGCTTGAAGCTTTGCGCAGCTCATCAACAAAGAGTTCCTGCTGAC 4806
Qy 4102 TTCAATCCGAGCTTGAAGCTTGAAGCTTTCATGAGGAGCGGCTGAGCAAGTGGCC 4161
Db 4807 TTCAATCCGAGCTTGAAGCTTGAAGCTTTCATGAGGAGCGGCTGAGCAAGTGGCC 4866
Qy 4162 TCATCATCATGACCGTGTGAGAGCAAGTGAAGTACGCTGATGTGCTGAAGCAG 4221

Db 4867 TCCTCATCATGACCGGCTGAGAGGCGGCTGGAATATGCACTGATGTCTTCAAGCAG 4926
Qy 4222 CTGCTGAGGAGCTCTTATGACAAAGCTGAGAGAGCAAGACCACTGAAGCTGCTGCTC 4281
Db 4927 CTGCTCTGAGCTCTATGATTAAGAACTGAGAGCAAGAAACCAAGGAGCTGATCTC 4986
Qy 4282 AGAGAGATGAGTCAATGCTGAGAGAAATGCTGAGCAATTTGATTTCTTCTCTTAC 4341
Db 4987 CGAGAGACAGAGTCTGTGCTGAGAGAAATGCTGAGCAATTTGATTTCTTCTCTGAC 5046
Qy 4342 AAGTTCCTAAGAGTGTGCTGAGAGGCGGCTCTTCTCTCTGTTCTGTGCTATGAGCAG 4401
Db 5047 AAGTTCCTAAGAGTGTGCTGAGAGGCGGCTCTTCTCTCTGTTCTGTGCTATGAGCAG 5106
Qy 4402 CAGATGAGAGAGGCGGCTTGAAGCCTATCAAGGCGGAGGCGGCTCTCTTCTGAGGAG 4461
Db 5107 CAGATGAGAGAGGCGGCTTGAAGCCTATCAAGGCGGAGGCGGCTCTCTCTGAGGAG 5166
Qy 4462 GACAACTCATCGCCAGAGATGATCAAAACCTGTGTCTGAGCTGTGAGCTCA 4521
Db 5167 GACAACTCATCGCCAGAGATGATCAAAACCTGTGTCTGAGCTGTGAGCTCA 5226
Qy 4522 GACAACTCATCGCCAGAGATGATCAAAACCTGTGTCTGAGCTGTGAGCTCA 4581
Db 5227 GACAACTCATCGCCAGAGATGATCAAAACCTGTGTCTGAGCTGTGAGCTCA 5286
Qy 4582 GTCAGAGAGAGATTTGAGTCCATCTTCAAGAAATGTGCTTCTGCTCCACCGGCGCAA 4641
Db 5287 GTCAGAGAGAGATTTGAGTCCATCTTCAAGAAATGTGCTTCTGCTCCACCGGCGCAA 5346
Qy 4642 GTCAGAGATGATGATGAGTGGGAGCAAGAAATGAGGCGGAGAGATGATTTGAGAT 4701
Db 5347 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5406
Qy 4702 GAAGATATCAACCAAGATTTGAATTTGAATTTGAATTTGAATTTGAATTTGAATTTGA 4761
Db 5407 GAAGATATCAACCAAGATTTGAATTTGAATTTGAATTTGAATTTGAATTTGAATTTGA 5466
Qy 4762 CAGATGAGAGATTTGAGTCCATCTTCAAGAAATGTGCTTCTGCTCCACCGGCGCAA 4821
Db 5467 CAGATGAGAGATTTGAGTCCATCTTCAAGAAATGTGCTTCTGCTCCACCGGCGCAA 5526
Qy 4822 GTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4881
Db 5527 CTTGCTCTGAGCAATCTTCCGAGCTCATGAGATGATGATGATGATGATGATGATGAT 5586
Qy 4882 ACCGAGAGCGGCGGAGCTTCTGAGTCAAGCACTGATGATGATGATGATGATGATGAT 4941
Db 5587 ACCGAGAGCGGCGGAGCTTCTGAGTCAAGCACTGATGATGATGATGATGATGATGAT 5646
Qy 4942 GAGATGAGATGAGTCAAGTGAAGAAACAGAGCAAGAGCAAGAGCAAGAGAGAGAGAG 5001
Db 5647 GAGATGAGATGAGTCAAGTGAAGAAACAGAGCAAGAGCAAGAGCAAGAGAGAGAGAG 5706
Qy 5002 CGGAGAGAGAGAGTGTGTAATTTCACTGATGATGATGATGATGATGATGATGATGAT 5061
Db 5707 CGGAGAGAGAGAGTGTGTAATTTCACTGATGATGATGATGATGATGATGATGATGAT 5766
Qy 5062 CTGAGAGAGTGTGATGATCTTTGAGACATCTTCAAGCAAGAGCAAGAGAGAGAGAG 5121
Db 5767 CTGAGAGAGTGTGATGATCTTTGAGACATCTTCAAGCAAGAGAGAGAGAGAGAGAG 5826
Qy 5122 GCGCTGCGGCTGAGCATCAAGTATGATGATGATGATGATGATGATGATGATGATGAT 5181
Db 5827 GCTCTCGGCTGAGCATCAAGTATGATGATGATGATGATGATGATGATGATGATGAT 5886
Qy 5182 GGCATTCATGAGCCGAGCTGCGCATCTGAGAGAGCAATTTGATGATGATGATGATGAT 5241
Db 5887 AGCATCATGAGCAAGATGAGGAGCACTGAGAGAGCAATTTGATGATGATGATGATGAT 5946
Qy 5242 TGGGTCATGATCAAGAAACCGGAGTGTGATGATGATGATGATGATGATGATGATGAT 5301
Db 5947 TGGGTCATGATCAAGAAACCGGAGTGTGATGATGATGATGATGATGATGATGATGAT 6006

QY 1048 CTGATAGAGTGGGCGGCTGTCATCTTCACTTGAAGCAGATTAATGACCGCATTAAGAG 1107
DB 1321 CCGGATGAGTCTGCGCTGTCGTCCTTCCCTATCCGGGCGCATCAACTTGAGATCAAGAG 1380
QY 1108 CCGCTGCACTCTTTTACCGGGGCGAGGCGACGCTGACCTGGCTGCTCAAGGTGAAG 1167
DB 1381 CCGCTGCACTCTGCTACCAAGGCGAGGCGACCTGAGCTCAACTGGCTGCTGGGGAAG 1440
QY 1168 GACATCCCTGACAGAGTGGCTCTTACATTAAGCATACTTCTGCTGGCTGACATG 1227
DB 1441 GACGTGCACTGACCAAGCGCGCTGTCCTCATGATGATTACTTGTGAGCTGACATC 1500
QY 1228 AATGCTGCGCTGGAGTGGCGACATGGTGGTGGAAATCCGCTTCACTGAGAGACAG 1287
DB 1501 AACGAGCTCTGGAGGCTCAACTCCAGTGAAGGCGCTGACCTGTATCACACACAGAG 1560
QY 1288 GACCGCATGACGTCTGTCAATGATGATGATCAAGAACCACTCTGGCTTGTGGG 1347
DB 1561 GACCGCATGACCTGTGTGCTCTCTTCAACGCGTACAGCGTGTGTTGTGGAG 1620
QY 1348 ACCAAAGTGGCAAGCTGAAGAGATCCGAGTGGATGACCCAGGAGCAAGCGCTCGAG 1407
DB 1621 ACTTAAGTGGCAAGCTGAAGAGATTCGAGGCGAGCGGTCCCGCATGCTGGGCTCGAG 1680
QY 1408 TATGAGACGCTGACAGTGT--GACCCCGGCTCAATCTCTGGGATATGCTTCTCC 1464
DB 1681 TACGATGATGCTCTGTGCTCAAGGACGAGACCCCATCTCGGAGCAATGGCTTCTCC 1740
QY 1465 AAGGACCAAGCAACTGATCATGTCAGAGAGGACGCTCAACAGAGCTCTGTGAG 1524
DB 1741 ATGATCAAGCGTACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
QY 1525 TCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1584
DB 1801 TCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
QY 1585 TCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1644
DB 1861 TCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
QY 1645 CCGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1704
DB 1921 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
QY 1705 TCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1764
DB 1981 TCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
QY 1765 GCTGCGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1824
DB 2041 GCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
QY 1825 CAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1884
DB 2101 CAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2157
QY 1885 GACGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1944
DB 2158 GACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2217
QY 1945 ACCAGCTTGTCTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2004
DB 2218 ACCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2277
QY 2005 CCAATACCGTGGCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2064
DB 2278 GCTTCCGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2337
QY 2065 TCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2124
DB 2338 TCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2397
QY 2125 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2184

DB 2398 GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2457
QY 2185 CAGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2244
DB 2458 CAGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2517
QY 2245 CAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2304
DB 2518 CAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2577
QY 2305 TATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2364
DB 2578 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2637
QY 2365 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2424
DB 2638 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2697
QY 2425 GAGAGTGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2484
DB 2698 GAGAGTGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2757
QY 2485 CAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2544
DB 2758 GAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2817
QY 2545 TCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2604
DB 2818 TCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2877
QY 2605 CCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2664
DB 2878 CCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2937
QY 2665 GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2724
DB 2938 GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2997
QY 2725 ATCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2784
DB 2998 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3057
QY 2785 TCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2844
DB 3058 CAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3117
QY 2845 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2904
DB 3118 TACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3177
QY 2905 GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2964
DB 3178 GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3237
QY 2965 TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3021
DB 3238 CTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3297
QY 3022 ACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3081
DB 3298 CCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3357
QY 3082 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3141
DB 3358 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3417
QY 3142 GAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3201
DB 3418 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3477
QY 3202 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3261

Db	3860	GTGAC	TACCGCC	CACAGT	CTGGAC	CACTGTGGAA	CGGCCAGATGAG	TTGGATTCTCTTTA	3919	
OY	3368	ACAAC	TCCAGT	TCCCTGCT	CATCTCA	CAAGACAA	CCAACTTCCATCTAC	TATCCAAACCGG	3427	
Db	3920	ACAAT	TTTAATCTCT	TACTAC	TCTATTA	ACGACAA	CGAATTCATCTCA	TCCCCAACCA	3979	
OY	3428	TGTTT	AGACCTTT	TGCTCC	CTCAGAA	TCTCGAGCT	CAAGCC	TGSCAGCCCATATCC	3487	
Db	3980	CGTTT	GACGTG	CTAGCC	CCCACTG	GAATCTTGAT	TCAGAA	CCAGGCTCACCATCATATCC	4039	
OY	3488	TAAAG	GGCAAGAA	CCGAT	TATCCG	CGCTGTGGTGGGGCA	CGTAGAC	CTGAACTACACG	3547	
Db	4040	TGAAG	GGCAAAAT	CTCTCT	CTCTGCT	CTGGAAG	---	GGCAACTTCACTACAG	4096	
OY	3548	TGCTG	TTGGGAGAA	CCGTG	CACCGCTG	ACCGTGTCA	ATGTCCAG	CTGTCTGCAAGT	3607	
Db	4097	TAA	TGATTTGAG	AGACAC	CTTGTA	CAAGCACTGTGT	CTGAGACAC	AGCTGCTTTGTGAC	4156	
OY	3608	CCCCA	ACCTCAT	TCGSCAG	GCACAA	AGATAG	TGCCCTGTAC	TGGTGCATGAGATCTCC	3667	
Db	4157	CTCC	CAACCTCAC	AGGGAC	GCACAA	AGTCA	AGTTTCA	CGTGGGGGAGTGTCTAC	4216	
OY	3668	CGGG	ATGGTAT	CAATG	CCCGGA	CACCCCTC	CACCTG	CCCCGCATCTGCAGTTCG	3722	
Db	4217	CTG	ACTCTCG	TGAGCGT	CATCTCC	GCACACCGT	TGAGCTTG	CCAGCATATCATGATTCG	4276	
OY	3728	CAGT	GTGTGCGG	CCCTCT	CTCATCAT	TTTTCAT	CGTGGCCG	TGCTCATTTGGCTTAAACGA	3787	
Db	4277	CAGCT	GTGGAAG	CCCTCT	TTATCAT	CGTCACTAT	TGTCTCAT	TGCTCATGCTTTCAAGCGCA	4336	
OY	3788	AGTCC	CGCGAAAGT	GCAC	CTCAG	CGTGAAG	CGAGTGC	AGATGAGGACCACTGGAGT	3844	
Db	4337	AGCT	TAGGAAAT	ATAGCT	CAACAT	CTCAAG	CGGCTTCA	AAATCGAAATCGATGAGT	4396	
OY	3848	CCCGT	GTGCGCT	CTGAG	ATGCA	AGAAAG	CCCTTGC	CGAGCTGCAGAGGACATTCATGAC	3907	
Db	4397	CAGG	GTGGGAC	CTGAG	TGCAAGAA	GGTTTGG	GAGCTT	CAGACAGCATCAATGAGC	4456	
OY	3908	TGAC	CAAGTGA	CTGAT	TGAG	CCGAGAT	TCCGTTCT	CTGAGACTATAGACCTTACACCATATGC	3967	
Db	4457	TAA	CAAGTGA	CTTGAT	CGATCAG	GAATCCCTT	CTAC	TGAGCTACCGTATGACATATG	4516	
OY	3968	GCGT	CGCTGTTCC	CGAAGAA	TTGAA	AGACCA	CCCTGTCC	CGGGA	CTTTGAGGTTCCGCGCT	4022
Db	4517	GAGT	CTCTTCC	CGACAT	TGAGAC	CAACCT	CTTCTG	TCGCGGA	CTGAGGATCAGGAA	4576
OY	4028	ACCG	GACAGAG	CGGTGT	GAGAAAG	GGCCTT	GAAAGCTTTG	CCAGCTCATCAACACAGG	4087	
Db	4577	ATGGA	CAGCGAC	CGTGGAGAA	AGCCCTT	GAACTTTG	CGCGAC	CTTATCAACACAGG	4638	
OY	4088	TGTT	CCGAGT	ACTT	CATCCG	CAGCTT	GAGTCC	CAAGCTATCTTCTCCATGCGGAC	4147	
Db	4637	TGTT	CTTGCG	GA	CTTCA	TCCGTA	CACTG	GAACATACAGCCGATCTTCCATGCGGAC	4696	
OY	4148	GTGG	CAACCG	GGCTCAT	CACTAC	AGACCG	GTGCGAG	AGACGCTGAGAGCTGAGATGCGGAC	4207	
Db	4697	GTGG	AAACG	TGGCTCT	CTCATAT	GACAGG	CTTCA	AGGGTCCGCTGAAATATGCCATG	4756	
OY	4208	ATGT	GTGAAC	CAAGCT	GTGCGG	ACCTCAT	TGACAA	GAACCTTGAAGACAAACACAC	4267	
Db	4757	ATGT	CTCAAG	ACAGCT	CTCTG	ACCTCAT	TGACAA	GAACCTTGAAGACAAACACAC	4816	
OY	4268	CTAG	CTGCTG	CTCAG	AGAGACT	GAGTGC	TGAGAA	ATGCTGACCAATTTGTTT	4322	
Db	4817	CCA	AGCTGCT	CTTCCG	CAGACT	GATGCT	GTGTG	CGAGAAATGCTGATCTAACTGGTTTG	4876	
OY	4328	CTTT	CTCTCT	CTACAG	ATTCCT	CAGAGAT	GTGCTG	GGAGACCCCTCTTCCCTGCTT	4387	
Db	4877	CTTT	CTCTCT	CTACAG	ATTCCT	CTGAAAG	ATGATG	CTGGGAAACCACTTCTATGCTTACT	4933	
OY	4388	GTGC	CACTCA	AGCAGAT	TAGAA	AGGGCC	CCCATTTG	ACGCTATCACGGGCGAGCCGCT	4447	

Db	493	GTGCATATCAGGAGCAGATGGAAAAAGGCCCATCTAGACCTATTCGTGGTAGGCCGAT	49936
Qy	4448	ACTCTCTTGAAGGAGGACAAAGCTCATCCGCGACAGATTTGACATACAAAACCCCTGCTCGA	4507
Db	4997	ACTCCCTGATGTAAMACAAAGTCTATCCGGGACAGATCGATGTAAAGATCTGATCTCGA	5056
Qy	4508	GCTGTGTACGCCCAACAATGCCAAACGCCCGAGGTCCCAATAAAGATCTCAACTGTG	4557
Db	5057	ACTGTGTCAACCTCTGACATATGAGAAAGCCCAAGATCCCAATGAAAGATCTAAACTGTG	5116
Qy	4568	AACACCATCACTAGGTCAAGGAGAAATTTGTGATGCATCTTCAGAAATGTGCCTTCT	4627
Db	5117	ACACCATCACTCAATCAAGAGAAAGATCTCGATGCCGATATATAGAAATGTCCCTACT	5126
Qy	4628	CCCAACGACCCCAAGCTGSCAGATTTGGATTTGGAGTGGGAGCAAGAAAGTGGGGCAGGA	4687
Db	5177	CCCAAGGGGCAAGGGCTGTGGACATGGATTTGAGTGGGCGCAAGCGGATTTGCCAAG	5236
Qy	4688	TGATCTTGGAGATGAAACATCAACCAACAAATGAGAAATGATTTGAAGCGACATGAACA	4747
Db	5237	TGAGTGTGAGGACGAAACATTTCCACCAAAATTAAGGATGATCTGGAAGCGGCTTTACA	5286
Qy	4748	CATGTGCCCCATACAGAGTGCACAAATGCTTCGATGGGCATTTAGTTCCAAGCAGGTGA	4807
Db	5297	CATGTATGATTTACAGAGTGTACACAGATTCGGTGTGCTGTGTCTTAAGCAGACCT	5356
Qy	4808	CAGCCTATTAACGCAATGAACAATCTCACCTCTTCAGACCTTCAGCAAGTAAATTTGAA	4867
Db	5357	CCCTCTACAAATCCCTGCTCTTCGACAGATCTCTGGACATCTATTGACAAATATGACT	5416
Qy	4868	ACATATCTCGGATACAGGGGACGCCCAACGCTCCGCTCAAGGACACCTATGATCACTC	4927
Db	5417	CTTCCTTACGATACACAGGACCCCAACAGCCTCCGATCCGGGTCCTCAAGATCAACC	5476
Qy	4928	CTGACCTGGAGATGGGATTCAGAAATGGGACCTTAATGAAAGCAACAGACACGGAGAC	4987
Db	5477	CAGACTTGAAGGCGGTGTCAAGGTTTGGCATCTGGTAAAGATCATGACATAGTGAACC	5536
Qy	4988	AGAAAGAGGGGAGCCGGGGAGCAAGATGTGTCTAAATCTTACTGACCCGACTCTGG	5047
Db	5537	AGAAAGAGGGTACCCGGGACAGAAATGTGTGATCTAATCTTGAACCCGGCTTCAG	5586
Qy	5048	CCACTAAGGGCACATGACAGAAATTTGTGATACCTCTTGAACCAATCTTCACACGG	5107
Db	5597	CCACCAAGGGCACCTTGCAAAAATTTGTGACGACTTTTGAACCTTGTTCAACATCG	5566
Qy	5108	CACACCGTGTCTTCCCTGCCCCCTGGCCCATGAATACATGTTGACTCTCCGATGAC	5167
Db	5657	TGCACCGGGATATGTCTTCCCCCTACCAATCAAGTACATGTTGATTTCTCGAATGACC	5716
Qy	5168	AGGCTGATTAACATGAGCATTCATGATCCCGGACGTCGCCCATACCTGGAGAGCAATTGCC	5227
Db	5717	AGGCAACACGACACAGTATTCACACACAGATGTGGGACACACCTGGAAAAACAATGCC	5776
Qy	5228	TGCCCCGATGTTTGGGTGCAATGATCAAGAACCCGACGTTTGTGATCATCATCA	5287
Db	5777	TTTCATTTGTCTTCCGGTGAATCTCATGAAGAACCTCAATTTGATTTGACATACACA	5836
Qy	5288	AGAAACGATCAACAACGCTGCTCTCTGTGTGCTCAAGACCTCAAGATCTTGTCT	5347
Db	5837	AGGGACGATCACAATATCTGCTCTCTGTGTGAAGCCAGACCTTATGAGACTCTGTT	5886
Qy	5348	CCAACGTACAGACACGGCTGGGACGAAGACTGCCTTCACCAAGCTGCTGTTATGCCAAG	5407
Db	5897	CCACATCAAGACACCGATGAGGACAGGACTCACCTTCACAAACAGCTGCTTATGCCAAG	5956
Qy	5408	ACATCCCCGATCAAGAAATTTGGGTGGAGAGTATTACTCAGCATATGGGAGAGATCCAG	5467
Db	5957	ATATCCCCGATTAATAGAACTGGGTTAAGAAATCTATGAGATATTTGCCAAGCTCCAG	6016
Qy	5468	CCATCAAGACCAAAACATGAACCATACTGTGCTGACAGTCCCGGATGCACATGAATG	5527
Db	6017	CCATTTGTACCAAAATATGAATCTCACTCCGAGGACAGTCCGCGCTGACTGCTTACAG	6076

QY 5528 AGTTCAACACCATGATGAGTCTCTGAGATCTTCTCTATGAGGCAATTAAGGAGG 5587
 DB 6077 AATTCAATATGCTGAGCCGCCCTCAACGAGATCTATATGTACGACATCACTGAGG 6136
 QY 5588 AGATCTTGAGACCTTGGACGACGATGACCACTGTGGAGACGAACTGGCTTACAAAC 5647
 DB 6137 AACTCATCGGGGCACTAGAGCAGATGAAAGAGCCGCAACGACGCTGGCTTACAGG 6196
 QY 5648 TAGAACAAGTCATAACCTCATGAGCTTGAACAGCTGAATTA 5691
 DB 6197 TGGAGATCTCATCAACGCACTGTCTCATAGAGACTGAAGGA 6240

RESULT 6
 ABT06294
 ID ABT06294 standard; cDNA; 5895 BP.
 XX ABT06294;
 AC
 XX
 DE 24-OCT-2002 (first entry)
 XX
 DE Human NOV12 coding sequence.
 XX
 KW Human; NOVX; autoimmune disease; cancer; infection; inflammatory disease;
 KW storage disorder; muscle disorder; neurodegenerative disorder; noctropic;
 KW developmental defect; neuroprotective; antiparkinsonian; hypotensive;
 KW hypertensive; haemostatic; cardiac; antidiabetic; dermatological;
 KW immunosuppressive; anti-inflammatory; virulence; antibacterial; anti-HIV;
 KW antiparasitic; antiallergic; antipneumatic; antineumatic; antineuritic;
 KW vulnarity; anorectic; antidiabetic; immunomodulator; antiparasitic;
 KW nephrotoxic; ketolytic; antidiabetic; cerebroprotective; anticonvulsant;
 KW antineuritic; antidiabetic; antidiabetic; antidiabetic; antidiabetic;
 KW tranquilizer; analgesic; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN MO200257450-A2.
 XX
 PD 25-JUL-2002.
 XX
 PF 29-NOV-2001; 2001MO-US048922.
 XX
 PR 29-NOV-2001; 2000US-0253834P.
 PR 30-NOV-2001; 2000US-0250926P.
 PR 25-JAN-2001; 2001US-0264180P.
 PR 20-AUG-2001; 2001US-0313656P.
 PR 05-OCT-2001; 2001US-0327456P.
 PR 28-NOV-2001; 2001US-00327456.
 XX
 PA (CURA-) CURAGEN CORP.
 PI Edinger S, MacDougall JR, Miller I, Ellerman K, Stone DJ;
 PI Gerlach V, Grose KW, Albrecht JP, Lepley DM, Rieger D, Burgess CE;
 PI Garmah SJ, Stryck K, Boldog FL, Li L, Padigan M, Mishra V;
 PI Baturajan M, Shenoy S, Raselli L, Tchernav VT, Verne CM;
 PI Zernsen BD, Malyankar UM, Guo X, Miller CE, Gangoli EA;
 XX
 DR WPI; 2002-590741/63.
 DR P-PSDB; AAO18749.
 XX
 PT Novel isolated polypeptide, designated NOVX, useful for treating or
 PT preventing in NOVX-associated disorders e.g. cardiomyopathy,
 PT atherosclerosis, diabetes, cancer, allergy, asthma, Crohn's disease.
 XX
 PS Claim 9; Page 130-131; 353pp; English.

XX
 SQ Sequence 5895 BP; 1172 A; 1928 C; 1759 G; 1036 T; 0 U; 0 Other;
 Query Match 44.5%; Score 2533.8; DB 6; Length 5895;
 Best Local Similarity 68.0%; Fred. No. 0;
 Matches 3812; Conservative 0; Mismatches 1712; Indels 81; Gaps 17;

QY 141 CGCCGAGGCTTTCAATACCTGAGTGTGATGAGAGACAGACACATTTACTTGGGGC 200
 DB 217 GAGTGGGGCTTACCACCTAGTGTGATGAGAGACAGAGAGAGTATGTGGGGC 276
 QY 201 CGTCAATCGGATTTAAGTCTCCAGCGACCTGAGAGCTTTGTGAGAGAGAG 260
 DB 277 AGTGAACCGCATTTAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 336
 QY 261 GCCGAGAGAGACACACCAAGTGTACCCAGCCGCACTGTGAGAGAGAGAG 320
 DB 337 CCTGTGAGAGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 396
 QY 321 CCTGACCAACCAACCAATGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 380
 DB 397 CTTGGAG 456
 QY 381 GATTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 440
 DB 457 GCTGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 516
 QY 441 CAAGCTGGGGAGAGCTTATCATAGAGAGAGAGAGAGAGAGAGAGAGAGAG 500
 DB 517 CAAGCTGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 576
 QY 501 CTGAGCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 554
 DB 577 CAGATGAG 636
 QY 555 TGCCAG 614
 DB 637 GGGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 696
 QY 615 CAAAGACTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 674
 DB 697 GGGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 756
 QY 675 GATGATTAAGATCTTCCGACACCTTCAACATCTTCACTTGAATTAATTA 734
 DB 757 AAGCTCAAGATCTTCCGACACCTTCAACATCTTGAATTAATTAATTA 816
 QY 735 CTATGCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 794
 DB 817 GTACAGCTTCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 876
 QY 795 TCACAG 854
 DB 877 GACTGAG 936
 QY 855 GAG 914
 DB 937 GAG 996
 QY 915 GAG 974
 DB 997 GAG 1056
 QY 975 CATTGAG 1034
 DB 1057 GCTGGGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1116
 QY 1035 GAAATGAATTCCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1094
 DB 1117 CCGAG 1176
 QY 1095 CCGAG 1154

Db	1177	GAAATTAAAGGAGGGCAATCCAGTCTCTGTACCGTGTGAAGGCAAGCTCTCCGCGGTG	1236
Oy	1155	GCTCAAGGTGAAGGACATCCCTCTGAGAGATGCGCTCTTAACTATGACATATCTTGTG	1214
Db	1237	GCTCTCAACAGGAGGCTGGGCTGTATCACTCGCCCTCTGACGATTCGATGACACTTCTG	1296
Oy	1215	TGGCCTGACATGATGCTGCCCTGGGAGTGTCCGACATGTGTGGGTGAATTCCCGTCTT	1274
Db	1297	CGGGGAGGACTTCAACGAGCCCTCTGGGGGACAGTCCACCATGTAGGGGGAAGCCCTGT	1356
Oy	1275	CACGGAGACAGGAGCCGCATGACCTCTGTCAATGCATATGTCTAACAAAGAACACTCTCT	1334
Db	1357	CGTGACAAAGATGATGTCCTTGACCGCGGTGGCTGCTTATACATATCGGGGCGGACTGT	1416
Oy	1335	GGCCTTTGTGGGACCAAAAGTGGCGACACTGAGAAGATCCGGGTGATGAGCCCA----	1390
Db	1417	GGTATTCGCCGGCAAGCGCAAGTGGCCGATCCGCAAGATCTCTGGGACCTCTCAAAACC	1476
Oy	1391	-----GGGGCAACGCCCTTCAGTATGAGACGTTGAGATGGGTGAGACCCCGGCCAGTCT	1445
Db	1477	CGTGTGGCGGCTCTGCTGGGCTTACGAGAGCGTGTGGCCAGAGGGGCAACCCCATCT	1536
Oy	1446	CCGGGATATGGCTTCTTCTCAAGGACCAAGACCACTTACATCATGTCAAGAGGCACT	1505
Db	1537	GCGAAGCTCGTCTCAGCGCCCAACCAAGATCCTTACCCCATGACCGGAAAGCAGGT	1596
Oy	1506	CACCAAGATCCCTGTGGAGTCTCTGTGTGATACAGATCAGAGCGCGCGAGTCCCTTGCTC	1555
Db	1597	GACGGGGGCTGTGGAGAGTGTGTGTGAGTACAGTCTTGTGAGCTGTGTGTGGGCTC	1656
Oy	1566	AAGCAGCCCCCACTGTGGCTGTGTGTGTGTGTGTGCAACAAGTGCACCCGGAAAGAGCGGTG	1625
Db	1657	ACGGGACCCCACTGTGGCTGTGTGTGTCTGTGACAGCATGTGTGTGCGCGGGGAGCCTGT	1716
Oy	1626	TGAGCGGTTCMAAGGAGCCCGGAGTTTGCTCCGAGATGAAAGCATGTGTCCGGCTGAC	1685
Db	1717	TGAGGAGAGAGAGGAGCCCGACGCTTGTGTGGGACCTGTGCAAGTGTGTGTGCACTGAC	1778
Oy	1686	GGTCCATCCCACAATATCTCCGTCTCTCAGTACAAAGTGC----	1742
Db	1777	TGTGAGCCCCCGCATGTGTGTGTGTGATCCATGTCCAGGTCCAGATCTTGTGTGCAAGC	1838
Oy	1743	GTACATATTCOCGGAGCTGTGACGTGGGCGTCACTGACCTTTTGAAGACCTGTGAGAT	1802
Db	1837	CTGGAACGTGCTTGACTTCAAGCTGTGGGTCAACTGTCTCTTCAAGGACTTCAAGGAATC	1899
Oy	1803	GGATGGGCTGTGTGTGTGGCAATCAGATCCAGTCTACTTCCCTGACGCCAAGAGGTGCC	1862
Db	1897	TGAGAGCGCTCTGAGAGATGTGCGGATCCATGCGCGCTCACCTCCGCGCGGAGGTGCG	1956
Oy	1863	CGGAGTCA-----TCACAGAAATGGGGGACCAACATGTGCTGACGTTCAAGTCAATTC	1918
Db	1957	GGCCCATCAGCGGGGGCAGAGGTGAGGGGAGACAGGGGGTGTGGAATCTCTACTTAAATC	2016
Oy	1917	AAAGAGACCGGCGATGACTTGTCCAGACCAAGCTTTGTCTTCTTAATGTGACGCTCA	1978
Db	2017	CAGGAGACAGGGAAAGATTGTGCTGTGTGACTTGTCTTCTTCAACTGTGACGGTCCA	2078
Oy	1977	CAATTG----TGCCTGTCTGTGTGGAAGTCAATACGGTGCACACTGGTGTAAATTAACG	2033
Db	2077	CCAGTTCGAGCTGTCTGTCTGTGTCAAGGCTCTTTTCCCTGTCACTGGTGAATTAACG	2133
Oy	2034	GCATGTCTGACCCCATGACCCCAAGACTGTCTCTTCCAGGAAGGCCGATGTAAGTGC	2099
Db	2137	CCAGGTGTGACACACAAAGTGTGTGATGCGCCTTCTGAGAGGGCCGTGTAAAGTGTG	2196
Oy	2094	CGAGGACTTCCCCCAAGTCTCTGAGTGAACAAGATCTGTGTGCGCCGTGTGAGGTATCA	2153
Db	2197	TGAGAGCTGTCCACAGATCTTCCCTCCACGCGAATCTTACGTGCCAGTGGAGTGTAA	2256
Oy	2154	GGCTTATCAGGCTGAAAGCCAAAGATCTTCCCAAGCCCAAGTCTTGGGAGAGGTGTACGA	2211
Db	2257	ACCCATCAGCTTGGCGCAGAACTGTGCACAGCCACAGTACAGGCAAGCGGTGATATGA	2316

QY	2211	ATGATCTCTCAAAATTCAAGGGGACGCGAGCGAGCGAGTGGCCCGCCCTGGCGCTTCAACAGCTC	2273
Db	2217	GTGCTCTTTCACATCCCGGGCAGCCCGGCCCTGTCCACCGCCCTGGCTTCAACAGCTC	2276
QY	2274	CAGGCTACAGTGGCCAGAACCTCTTATTTCTATGAAAGGATGAGATCAACAACCTGCC	2333
Db	2377	CAGGCTCAGAGGCCAGAAATTCCTGTACTCTTCAGAGGGAAGATGTCAAGCACTGCC	2436
QY	2334	CGTGAAGTTGACAGTGGTGTGGAATGGGACCTTCAACATTGACCAACCACCTCAGAAATAA	2393
Db	2437	AGTGAACCTGTCACTGTGTGGAAACGGCAACTTTGTTCATTGACAAACCAAGAAACATCCA	2496
QY	2394	AGTTCACTCTTACAAAGTGTGAAGCCATGGCTGAAGACTGCGGGCTGTGCTTCAAGCTGA	2453
Db	2497	GGCGCACTCTTACAAAGTGGCCCGGCCCTGGCCGAGAGCTGGCGCTTCTGCTCAAGGCCGA	2556
QY	2454	CCCAAGACTTGCAATGTGGCTGGTGGACGAGGCCCAAGGCAAGTGAACCTGTGGCCGACACTG	2513
Db	2557	CCGCGCTTGAAGTGGGAATGGTGGCTGGTGGCCGAGCGCGCTGCTCTTGTGGACACCACTG	2616
QY	2514	CCCTGCCCA--GGAGAGCCAGTGGCTGAAGCTGTCTGTGTCCAAAAGCAAGTGCACAA	2570
Db	2617	CGCTGCCGACACACTGCATCTGTGAATGACAGCGCGTCAAGGACAGTGGCTGCACCGA	2676
QY	2571	CCCCCGCATCAGAGATTAATCCCGGTGACAGCGCCCCCGGAAAGGGGGGACCCAAAGTCAAC	2630
Db	2677	CCCCAAGATCTTCAAGCTGTCCCGGAGAGGGGCCCAAGGCAAGGCGGCAAGCGGCTCAC	2736
QY	2631	TATCCGAGGGGAGAACCTGGGGCTTGAATTTCCGCAATCGCCTCCCATGTCAAGTTGC	2690
Db	2737	TATCAGAGGGAGAACCTGGGGCTTGGCATTCGAAAGACGTGCTGTGGGCGCGCGTGGG	2796
QY	2691	TGGCGTGAAGTGAAGCCCTTTAGTGAATGCTTTCATCCCTGCGAACAAGATCGTGTGTA	2750
Db	2797	CAAGTGTGTGACAGCCCTGTGGAGAGCGAGTACATAGTGGGAGCAAGATCGTGTGTA	2856
QY	2751	GATGGGGAGAGCCCAAGCCC--AGCCAGCATGACAGCTTGTGTGAAGATCTGTGTTGCTGT	2807
Db	2857	GATGGGGAGAGCCCAAGCTCCGTTGCGTGGCCATGACGCCCTGGTGAAGGTGTGTGGGGGA	2916
QY	2808	GTGTCCGCTGAATTCAATGAGCCCGGTCTCAAGCTCTATTACTTATGACATGACTCT	2867
Db	2917	CTGTCTACACACACTACCGCGCCCTGTCAACCAAGCTTCACTTGTGTGAACCAACTT	2976
QY	2868	CTCAGATCTGAAGCCACAGCGGGGGCCCATGTCCGAGGGGACCCAAAGTGAACATCAAGG	2927
Db	2977	CTACCGTGTAGCCCTCCCTCCGTGTGGCCCTGTGTCAAGGGGGGACCTGGAATGGCATGAGGG	3036
QY	2928	CACCAACTGAATGCCCGAGACCACTGTGTGTATTTTGGAAAAGACCCCTGTCTT	2987
Db	3037	AAGCACTTGAACCGAGCACTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTCTT	3096
QY	2988	CCACAGGCGAATCTCATCTCTTCAATGTCTGTGACACACACATCTCAGATGAGTGTAGA	3047
Db	3097	CTCTGTGTCCAGAGGAACCTCCGTGAAGATCGGTGTGCTGAACCCCCCGGGCAGAGCC	3156
QY	3048	GATGAAGTGTGGTGAAGTGTGAAGAGCC-----AAGATCCACCGAGACCTGTGT	3098
Db	3157	TGGCAGCCCTCCCATCATCATCAACATCAACCGGCCCAAGTCAACAACCTTGAAGTGA	3216
QY	3099	CTTTCAGTATGTGAAGAAGCCCAACATCGTGGCATTTGAGCCAAATGAGACATTTGTCAG	3158
Db	3217	GTACAACTAACCGAGACCCCAACATCTGAGATGACCCCGAGTGAACATCAACAG	3276
QY	3159	TGAAACACACCCATCGCCCTTATGAGGGGACCCACCTGACCTCATPACGAAGCCCCAGAT	3218
Db	3277	CGGTGGAGCCCTCTGAAGTCAACAGGACCAACCTGCACTGTGCGTGAACCCCAAT	3338
QY	3219	CGGTGCCAAGCATGAGGAGGAAGAGCAACATCATCTGTGAGGTTTGAACGCTACTGA	3278
Db	3337	CGGGCCAAAGTATGAGGCAATTGAG---AGGAGAACTGCTGTGTATCATGACACAC	3393

Accession	Gene	Protein	Location/Qualifiers	RefSeq ID
D8	5536	GGTGGAGAGGAGGTA	CTACTATGACAGACATGCCAAATGTCGACCCATCAGCCACAGAGACAT	5593
OY	5487	GAACGCATACCTGGCTG	TGACAGCTCCGGATGCACTGATATGATCTCAACACCATGATGTC	5546
D8	5596	GAGTGGCGATCTGGCTG	TGACAGATCCCGCTTCACCTGAGCCAGATTCACACATGAGCGC	5555
OY	5547	ACTCTCAGAGATCTCTCT	CTTATGTTGGGCAATAACAGCGCA-----GGAGATCCTTTGAC	5600
D8	5656	CTTGACACAGATCTACT	CTCTCTACATACCAAGTACAGAGATGAGTGGACGATCTTGACAGC	5715
OY	5601	TCTGACACACATGACCA	CGTGTGGAGACAGAACTGGCTCTTACAACTAGAACATGTCAT	5660
D8	5716	CCTGAGAGAGATGTGAC	GCGCGCGGACAGCGCTGCGAGACAAAGCTGAGACAGTGGT	5775
OY	5661	AACCTCATGAGCTTAC	AGAGCTGA 5685	
D8	5776	GGACACGATGGCCCT	GAGACGCTGA 5800	
RESULT 7				
ADAO9914	ID	ADAO9914	standard; CDNA; 6826 BP.	
AC	ADA09914;			
XX	20-NOV-2003	(first entry)		
DE	Human receptor and membrane-associated protein REMP-18	CDNA, SEQ ID:56.		
XX	Human; receptor and membrane-associated protein; REMP;			
KW	cell proliferative disorder; cancer; autoimmune disorder;			
KW	inflammatory disorder; infection; neurological disorder;			
KW	metabolic disorder; developmental disorder; endocrine disorder;			
KW	cytostatic; immunosuppressive; antiinflammatory; neuroprotective;			
XX	nocrotic; cerebroprotective; gene therapy; gene; ss.			
OS	Homo sapiens.			
XX	Key	Location/Qualifiers		
FT	CDS	176..2857		
FT		/tag="a		
FT		/product="REMP-18"		
XX	MO2003070902-A2.			
XX	28-AUG-2003.			
PD	18-FEB-2003; 2003WO-US004902.			
XX	20-FEB-2002; 2002US-0358279P.			
PR	13-MAR-2002; 2002US-0364338P.			
PR	25-APR-2002; 2002US-0375637P.			
PR	29-APR-2002; 2002US-0376699P.			
PR	10-MAY-2002; 2002US-0379837P.			
XX	10-MAY-2002; 2002US-0379853P.			
PA	(INCY-) INCYTE GENOMICS INC.			
XX	Chavala NK, Yue H, Richardson TW, Margulis JP, Lehr-Mason PM;			
PI	Gorvaid AE, Becha SD, Kable AE, Swarnakar A, Jin P, Hawkins PR;			
PI	Chien D, Ramakumar U, Tran UK, Hafalia AUA, Baugan MR, Lee SY;			
PI	Jiang X, Jackson AA, Khare R, Bulloch SA;			
DR	WPI; 2003-697610/66.			
DR	P-PSDB; ADA09876.			
XX	New human receptor and membrane associated proteins and nucleic acids,			
PT	useful for diagnosing, treating or preventing e.g. viral, bacterial,			
PT	fungal, parasitic, protozoan or helminthic infections, cancers,			
PT	neurological disorders.			
PS	Claim 5, Page 279-281, 298pp; English.			
XX				

Query Match	Best Local Similarity	Score	DB	Length
Matches 3730; Conservative 0; Mismatches 1777; Indels 108; Gaps 10;	43.2%;	2455.8;	8;	6826;
152	TCATTCACCTGGTGGTGGATGAGAGACAGAGACATTTTCTGGGGGCGGTCAATCGGA	211		
273	TTATCCACCTGGTGGTGGATGAGAGACAGAGACATTTTCTGGGGGCGGTCAATCGGA	332		
212	TTTAAAGAGCTTCCAGAGCCTTGAAGCTTTTGGTGAAGCATGAGACAGGGCGGACGAG	271		
333	TCCTTAAAGCTGGCCCCCAACTGACTGAGCTGGGGGCCCATGTCACGGGGCGCTCGAG	392		
272	ACAACCCCAAGTGTACCCACCCCGCATGCGCAGACCTGCAATGAGCCCTTGACACCA	331		
393	ACAAGCGCTGGTACCCCGCCCGCCCGCATGAGCGGTGGTGGCCACCGCTGGCCCCG	452		
332	CCAAATGTCAACAAGTGTCTCTCATAGCTTCAAGAGAAACAAGGTGATTTGGCTGTG	391		
453	TGGACAAATCAACAAGTGTCTCTCATAGCTTCAAGAGAAACAAGGTGATTTGGCTGTG	512		
392	GGAGCGGTGACCAAGGATCTGCAAGTGTGCAAGTGTGAGAGACCTCTTCAAGCTGGGG	451		
513	GCAGATGTGGCAGGGGATCTGCCAGTTCCTGGCTGGAAGACCTCTTCAAGCTGGGG	572		
452	AGCTTTATCAATAAGAGAGATCTCTGTCAAGTGTGCAAGAGAGGGCGTCAAGCTTTG	511		
573	AGCGGACCAACCGAAGAGAGATCTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	632		
633	GTGTATTTGTGAG	692		
512	GAGTATGTGTCTCTCAAG	571		
632	ATGCGATGTGTCTCTCAAG	691		
753	CGGATATGTGTCTCTCAAG	812		
632	CGGATATGTGTCTCTCAAG	691		
813	CAG	872		
752	GCAGATGTGTCTCTCAAG	811		

Db 873 CCTCCTTGGTATCTTCCGTGACCGCTGACGTGAGACCCAGACAGACGCTTTGACACAG 932
Qy 812 CCAAGAGCAGAGGTATATACATCCAGCTGTGAGGCTTTGCAAGAGACACAGCCTTCA 871
Db 933 CGGCGCAAGAAATTTTTCAGCTCCAGATGCTGCGCATGTGCGGGAGACTGAGATTCT 992
Qy 872 ACTCCATATGAGAGGTGGCCATTTGGCTGTGAGCGAGTGGGGGTGAGTACCGCTGCTGC 931
Db 993 ACTCATACGTGGAATTCCTCCATCGCTGCTCCTGAGCGCGGCTGAGTACCGCTTGGTGC 1052
Qy 932 AGGCTGCTTACCTGTCCAAAGCGGGGCGGTGTTGGCAGAGCCCTTGGATTCATCCAG 991
Db 1053 AGAGCGCCACCTGGCCCAAGCCTGGCTGTGCTGAGCCAGAGCCTTGGGCGTGCAGCTG 1112
Qy 992 ATGATACCTGCTCTTCAACGCTTCTCCAGAGGCCAGAAAGGAGAAATGAAATCCCTGG 1051
Db 1113 ATGAGAGCGTCTCTTCAACCATTTCTCTAGGAGCCAGAAAGACGGGCGCACCCCA 1172
Qy 1052 ATGAGTGGGCGCTGTGACATCTTCAATGAGCAGATTAATGACCGCATTTAAGACGGC 1111
Db 1173 GGGAGACATCTCTGTGCTTCAACCTCAGCAACATCAATGCGCATCCGCGCGCGCA 1232
Qy 1112 TGCAGTCTTGTACCGGGGCGAGGGGACGCTGGACCTGGCTGGCTCAAGGTGAAGACA 1171
Db 1233 TCCAGTCTGTCTATGTGGGAGGGGACCTGTGCTGCTGGCTGTGAAACAAGAGAC 1292
Qy 1172 TCCCTGACAGCAGTGGGCTTTAACCATTGACAGATTAATTCTGTGGCTTGAATG 1231
Db 1293 TGGCTGACATCAACCCCATGAGATCAAGGGCACTTCTGTGGGCGGTGTTGAACC 1352
Qy 1232 CTCCCTCGGAGTGTCCGACATGTGTGCTGGAATTCCTGCTTTACGAGAGACAGGACC 1291
Db 1353 AGCCTGTGGAGGCGCTGACATGTGATGAGGGGCTGCGCTGTGCGGCAACAGACCGAG 1412
Qy 1292 GCATGACGTGTCTATGCGCATATGTCTACAGAAACCACTCTGTGGCTTTGAGGACCA 1351
Db 1413 GCATGGCCAGCGTGGCGGCTTACCTTACCGCAGACCTGTGTGCTTCACTTGGCAGCG 1472
Qy 1352 AAAGTGCAGAGCTGAAGAAAGATCCGGGTGATGAGACCCAGGGGCAACGCTTCCAGTATG 1411
Db 1473 GCGAGCGGAGCTTGAAGAAAGTGGGGGTGATGCTTCCAGG--ATGGCCACCTGTATG 1529
Qy 1412 AGACGGGAGGTGGTGGAGCCCGGGCCAGTCTCCGGGATATGCGCTTCCAGAGACC 1471
Db 1530 AGACAGTCCCGGTGGTGGATGGAGCCCATCTCCGAACTGTCTTCAAGCCGGACC 1589
Qy 1472 ACAGCAACTCTTACATATGTCAAGAGGAGCTCAACAGATCCCTGTGAGTCTGTG 1531
Db 1590 ACCGGGACATCTATCTCTAGTGAAGAGAGGTGAGCCAGCTCCGGGTGAGACCTGTG 1649
Qy 1532 GTCAATATGAGACTGGGCGAGTGGCTTGGCTCAAGCGACCCCGCATGTGTGCTGTG 1591
Db 1650 AGCAGTACCAAGAGCTGGCAGCTGCTGGCTTCGGGAGCCCGCATGTGTGTGTG 1709
Qy 1592 TGGTGCACAACAGTGCACCCGAGAGAGCGGTGTGAGCGGTCAAGAGACCCCGAGGT 1651
Db 1710 TGGCTGGAGACAGAGTGTGCTCGGAGAGGGGCTGTCTGGGGCGCTGTGCCCCACAGGCT 1769
Qy 1652 TTGGCTTGGAGATGAAGAGATGTGTCCGGCTGACGCTCAATCCCAACATATCTCGTCT 1711
Db 1770 TTGCTGAGGAGCTGAGCAAGTGTGTCCAGGTGGGGTCCGGCCCAACATATGTCACTGA 1829
Qy 1712 CTCAATACACAGTGTGCTGTCTGTGAGAGCTTACATGTCCGAGCTGTGAGCTGCG 1771
Db 1830 CGTCACTGGGGGTGAGAGCTTACGCTGACCTGCAACAGTGGCCAGACCTCAAGTGGGGG 1889
Qy 1772 TCAACTGACACTTTGAGAGCCTGTCAAGATGAGATGG--CTGTGTGTGGGCAATCAGA 1828
Db 1890 TGAAGTGGCGCTTGAAGCGGCGGAGAAAGAGCGGTCTGTGCGCTCCGGTGAAC 1949
Qy 1829 TCCAGTCTACTCCCTGACAGCCAAAGAGGTGCCCCGATCATCAGAGAAATGGGAGACC 1888

Db 1950 TGTCTGCCCCCTACCTCCCTCCAGGAGCTCCAGCTTTCACAGGGGGAGTGGGGCCA 2009
Qy 1889 ACCATGTCTACAGCTTTCAGCTCAAAATCAAGAGACCCGCAATGACCTTCCAGACCA 1948
Db 2010 CCGGCACTGTGCGGCTGACAGCTTCTCTCAAGAGAGACGGGTGAGGTTGGCGGTGCTG 2069
Qy 1949 GCTTGTCTTCTCAATATGACGCTCCACAAATTCGTGCTGTGCTGTGAGAGTCCAT 2008
Db 2070 ACTTGTCTTCAACTGACAGGCTCTCCAGTGTGAGCATGTCTGTGTGGACGCTT 2129
Qy 2009 ACCGCTGCACTGTGTAAATACCGGATGTCTGACACCATGACCCCAAGACCTGCTCT 2068
Db 2130 ACCCTGCGCACTGTGTAAATCCGCAACGTTGACAGCCGCCCCCAAGATGCTCTT 2189
Qy 2069 TCCAGAAAGCCAGTGAAGCTGCGCCAGAGACTGCCCCAGCTGTGCTGAGTGAACAGA 2128
Db 2190 TCCAGAGGGGAGGCTCCAGACGCTTGAAGGCTGCTCCCTGAGATCTGTCCAGTGGGAGCC 2249
Qy 2129 TCTGTGTGCGGTGAGGTGATCAAGCTTATCAGCTGAAGGCCAAGAACTTCCCTAGC 2188
Db 2250 TCTGATCCCGTGTGGGTGATGACGCTCTTACTTTCGCGGCTAAGAACTTACTCAGC 2309
Qy 2189 CCGAGTGGGCGCGTGGCTAAGATGCAATCCCTCAATTCAAGGCGAGGAGAGAG 2248
Db 2310 CGCAGTGGGCGAGAAACTATGATGTGCTGTGCGGGTGCAGGGGCGGACAGCGGG 2369
Qy 2249 TGGCGGCTGTGCGCTTCAACAGCTCCAGCGTACAGTGCAGAAACCTTTATTTCTATG 2308
Db 2370 TGGCTGCGGTGCTTCAACAGCAGAGTGTGAGTCCAGAAAGCTCTTACTCTATG 2429
Qy 2309 AAGGATGAGATCAACAACCTGCGGTGAGTTGACAGTGTGTGAATGGGCACTTCA 2368
Db 2430 AAGGTATGACATGTGACACCGAGCTGAACTTCTCGTGTGTGGAATGGAATCTTCC 2489
Qy 2369 ACATTTGACAAACCACTCAGATTAAGTTCACTTCAAAAGTGTGAGACATGCTGAGA 2428
Db 2490 CCAATGACAAAGCTCCAGGCTCCAGACCTCTGTGCAAGTGTGCGGCGCAGGGGCCA 2549
Qy 2429 GCTGCGGCTGTGCTTCAAGCTGACCCAGACCTTCCGATGTGTGCTGCGGCGCCAG 2488
Db 2550 GCTGTGCTGTGCTTCAAGCTGATCCCGCTTCACTGTGTGTGATGATCTCAGAGC 2609
Qy 2489 GCGAGTGCACCTTGCAGACCTGCGCTCCAGAGAGACCAAGTGTGAGCTGATCTG 2548
Db 2610 ACAGTGCAGAGTGCAGACCCAGCTGCGGCGCCGAGAACCACTGTGATCACTGAGCC 2669
Qy 2549 GTGCCAAAAGCAATGACAAACCCCGCATCACAGAGATATCCGGTGAACAGCCCCC 2608
Db 2670 AGAAGGACACCGGTGAGCCACCTCCGATCACGAGATCACCTCTGTGGGGGCCA 2729
Qy 2609 GGAAGGGGGGCAACAAGTCACTATCCAGGGGAGAACTGGGCTTGAATTTGCGACA 2668
Db 2730 AGAAGAGAGGACCCGGGTCAACATGTGTGTGACAACTGGGCTTGTGCCGAGAGG 2789
Qy 2669 TCGCTCCCATGTCAAGTGTGTGCGGTGAGTGCAGCCCTTATGTGATGTATCATCC 2728
Db 2790 TGGGCTCTGCGGGTGTGCGGTGTGCACTCCATTCGCGGCGAGTATCA 2843
Qy 2729 CTGCAAG----- 2735
Db 2844 GTGCTGAAGGTGAGTGGGCTGTGTGGTGGCCCGGCGGTATGTGGCTGCGCGGCTT 2903
Qy 2736 -----ACAGTGTGTGAGATGGGGAGGC--CAAGCCAGCCAGC 2776
Db 2904 GACGCTCTGAGCCCTAGATCTGTGTGAGATGAGAGAGTCCCTGTGATGCCAGCCGC 2963
Qy 2777 ATGAGGCTTGTGGAGATCTGCTGTGTGTGCTGCTGAGCTGAATTCATGCGCCGCTCT 2836
Db 2964 CGCGGGGCGCGTGAAGT 3023
Qy 2837 CACAGCTTATTTACTTATGACATGACTCTTTCAGATCTTGAAGCCAGCGGGGGGCCA 2896
Db 3024 AGCAGGTCTACAGCTTGTGACCCCAAGTGTGACCAAGTGAAGTCCAGCGCGGCGG 3083

QY 2897 TGTCCGAGGAGGACCAAGTACATTCAGAGGACCAACCTGTAATCCGGAAGCAACGTGG 2956
Db 3084 CGTCCGAGGAGGACCAAGGCTTACATTCAGAGGAGCTCTGGATGCTGGCAGAGGGGTCA 3143
QY 2957 TGGTATGTTTGAAGAAGCAGCCCTGTCTCTTCCACAGGCGATCTCATCTCACTATGCT 3016
Db 3144 CAGTACCTGTGAGGAGCAGCGAGTCCAGTTTGTAAAGAGATGCCAAAGGCGATCGGT 3203
QY 3017 GCACACACCATCTCTCAGATGAGTGTAG--AGATGAAGGTGCGTGTGAGGTGACA 3073
Db 3204 GCATCTCAGCTCTCTCCACCCCTGGGCCAGGACGCCCCCATCACTTGGCATTTAGCC 3263
QY 3074 GGGCCAGATCCA--CAGGACTGTGTCTTCAGTATGTGAAGACCCACCATCGTGC 3130
Db 3264 GGGCTAACATCTCCAGCCCCGGGCTCATCTACCTACATCAGAGACCCACCGTCAACC 3323
QY 3131 GGATTGAGCGAGATGAGACATTTGTCAGTGAACACACCCATGCGGTATGGGAGACC 3190
Db 3324 GCCTTGAAGCCACCTGAGACATCATATGAGAGACGTCCATCACTGTGAGTGGAGCC 3383
QY 3191 ACCTGACCTCATACAGAACCCCGAGATCCGTGCGCAGAGATGAGAGGAGAGACATCA 3250
Db 3384 ACCTGCTAGCGTCCAGAGACCCCGGGTCCGTGCGAAGTACCGGCGCATGAGACACCA 3443
QY 3251 AATCTGTAGAGTTTGAACCGCTACCTGAGATGACCTGTCAAGCGCGCCGCTCGTCTGG 3310
Db 3444 ATACATGCCAAGTATCAAGACACTGCGATCTGTGAAGCCCCCGCATCTTTCTTG 3503
QY 3311 GTCTGACCAACAGTCAAGACTGACCGAGAGGCCGAGAGATTGTGCTTCACTCTGACA 3370
Db 3504 GGGGCGCCAGCCTCGGCGGAGAGGACACCTGATGATGTTGGTTCCTGCTGAGCC 3563
QY 3371 AGGTCCAGTCCGTCTCATCTCTCAACAAACCACTTCACTCACTCACTCACTCACT 3430
Db 3564 AGGTGAAACGGCCCGCTCTCAACCGCTCTCTTACCTTACCTTACCTTACCTTACCT 3623
QY 3431 TTGAGGCTTTGGTCCCTCAGAGATCTTGAAGTCAAGCTGAGCCGCACTCATCTTAA 3490
Db 3624 TTGAGCGGCTGGGGGCTCTGGCGTGTGGAAGTCAACCGGGCTCCACGAGGTGCTGA 3683
QY 3491 AAGGCAAGAACTGATCCCGCTGTGTGCTGGGGGCAAGTGAAGTGAACACTGTGC 3550
Db 3684 AAGGCAAGAACTGAT--TCCCGGGCAGCGGAGCTCCGCTCAACTACATGAGC 3740
QY 3551 TGGTGGGAGAGAGCGGTGCAACGTGACCGTGTCAAGTGTCCAGCTGCTGCGAGTCC 3610
Db 3741 TGTATGAGAGGCAAGCGGTGTGTGCTCACTGTCTCGGACACAACTCTGTGCGACTAC 3800
QY 3611 CCACCTCATCGCAGGACCAAGATGATGCGCGTGTGGTGGCATGAGTACTCCCGG 3670
Db 3801 CAGCCAGAGCTGGCCGCGAGCGCTGTATGATCTGGTGGTGGCGCTGAGATGTGGCTGG 3860
QY 3671 GAGTGTGTATCATTTGCCCGGACAGCGCGTCAAGCTGCGCCGCACTGAGCATCGAG 3730
Db 3861 GCACTTGTACATCTGGCAGAGCGGCGCTACCTTACCGGCACTGATGGGGCTGGGG 3920
QY 3731 TGGCTGGCGGCTCTCATCATTTTCATGATGAGCGGTGCTCATTTAAACGCAAGT 3790
Db 3921 CGGGGGGTGGGCTCTGTGTGTGGCATACAGCGTGTGGTGGCGTGAACGCGAAGA 3980
QY 3791 CCCGCAAGATGACCTCAGCTGAGAGCGGCTGACATGAGATGAGACAACTGAGATCC 3850
Db 3981 CTGAGAGCGGAGCGCTACCTTCAGCGCTGTGAGTGAAGTGAAGCAACTGAGATCC 4040
QY 3851 GTGTGGCCCTGAGATGACAGAGAGCTTTGCCAGAGTGAAGCGACATTCATGAGCTGA 3910
Db 4041 GTGTGGCCCTGAGATGACAGAGAGCTTTGCAAGAGTGAAGCGACATTCATGAGCTGA 4100
QY 3911 CCAAGTACCTGAGTGAAGCGGAGATTCGCTTCTGAGATTAAGAACTTAACCAAGCGGG 3970
Db 4101 CTAAACCAATGAGCGAGTGAAGATCCCTTCTGTGACTACCGGACTTACGCTGCGCG 4160

QY 3971 TGGTGTTCGAGGAATTGAAGACCACTGTCTCTCGGAGCTTGAAGTCCGGGCTTACC 4030
Db 4161 TGGTCTTCCCGGAGCATCGAGGCGCACCCGGTCTCAAGGAGCTGAATACGCCCAAC- 4219
QY 4031 GCGAGAGAGGTGTGAGGAAGGCTTGAAGCTCTTGGCCAGGCTCATCAACAAAGGTGT 4090
Db 4220 -----GTGAAGAGGCGCTGGCGCTCTTGGGAGAGTGTGTGCAACCGCGGT 4268
QY 4091 TCTGTGTCTCTTCAATCGCAGCTTGAATGCCAGGTAGTCTTCCATGCGGACCGTG 4150
Db 4269 TCGTGTACTCTTCACTCACAGCTGAGAGCGCCAGAGCGCTTCTCATGCGCGACCGG 4328
QY 4151 GCAAGTGGCTCACTCATATGACCGGTCTCAGAGCAAGCTGGAGTACGCGCATGTAG 4210
Db 4329 GCAAGTGGCTGTCTCAACATGTGTGGCGCTGAGAGCGGCTCACTATGCGACGAGC 4388
QY 4211 TGGTGAAGAGCTGTGCGCGACCTTCAATGACAGAACCTGAGAGCAAGAACCTTA 4270
Db 4389 TGGTCAAGCAACTGTGGCGGACCTTCACTGAGAGAACCTGAGAGCAAGAACCTCA 4448
QY 4271 AGCTGTGTCTCAGAGAGACTGAGTCAAGTGTGAGAGATGCTGAAGTGTGATTTACT 4330
Db 4449 AGCTGTGTCTCAGAGAGAGAGTCAAGTGTGAGAGATGCTTACCACTGCTTCACT 4508
QY 4331 TCTCTCTTCAAGATTCTCAAGAGTGTGTGGGAGCGCCCTTCTCTCTGTTCTGTG 4390
Db 4509 TCTGTGTGACAAAGTTTCTGAAGAGTGTGTGGGAGCGCTCTTCTCTGTTACTGTG 4568
QY 4391 CCAATCAAGCAGAGATGAGAGAGGCGCCCATTAAGCTCATACGGGCGAGGCGGCTACT 4450
Db 4569 CCAATCAAGCAGAGATGAGAGAGGCGCCCATTAAGCTCATACGGGCGAGGCGGCTACT 4628
QY 4451 CTTGAGAGAGAGCAAGCTCATCCGCGCAGAGATTGACTTCAAAACCTTGTCTGAGCT 4510
Db 4629 CCTGAGAGAGAGCAAGCTCATCCGTCAGAGATGAGTCAAGACATGACCTTCACT 4688
QY 4511 GTGTGAGCCGAGCAATTCGCAACAGCGCGAGGTCCGATTAAGATCTCACTGTGANA 4570
Db 4689 GCGTGTGTCCGAGAGAGAGGCGAGCGCCAGGTCCGATGAGAGTCTCAACTGTGACA 4748
QY 4571 CCACTGAGGTCAAGAGAGAGATTTCTGATGCCATCTTCAAGAAATGTGCTGTCC 4630
Db 4749 GCACTACCGAGGCAAGATTAAGTCTGTGACATGTGTACAAAGGCACTTCGTACTCC 4808
QY 4631 ACCGGCCAAAGCTGCAATATGATGTGAGTGTGCGACCAAGAGAGTGGGAGATGA 4690
Db 4809 AGCGTCCAAAGCTGAGGACATGAGCTGAGTGGCGGCGCAGGCGCATGACTCGCATCA 4868
QY 4691 TCTTGAAGATGAGAACATCAACCAAGATGAGAAATGATGAGAGGCACTGAACAC 4750
Db 4869 TCTTGAAGATGAGATGTACCAACCAAGTGAAGTGAATGAGAGGCTCACTAC 4928
QY 4751 TGGCCCACTACCAAGGTGCAAGTGTGCTGTGAGCAATTAAGTTCAAAGAGTGAAG 4810
Db 4929 TGGCCCACTACCAAGGTGCAAGAGTGTCTGTGAGCAATTAAGTTCAAAGAGTGAAG 4988
QY 4811 CCAATTAAGAGTGAACCACTCAACGCTTCCAGAGCTCAGCAAGTAAATTAAGAAACA 4870
Db 4989 CCAATTAAGAGTGAACCACTCTTCACTTCAACCG--CTTCTCAAGCGCTACAGAGCT 5045
QY 4871 TGAATCGGATCAAGGAGAGCGCCGACAGGCTCCGCTCAAGAGCACTATGATCACTCTG 4930
Db 5046 TGTCTCGCAAGGCAAGAGCGCTGATAGGCTCGCTCAAGGCGACCAATGATTAACGCTG 5105
QY 4931 ACCTGAGAGTGAAGTCAAGATGTGGCACTTGAAGAGCAAGAGCGAGCGAGAGCAAGA 4990
Db 5106 ACCAGAGAGAGGACCAATTTGTGACCTGTGAGAAACCAAGAGCACTGCGGACATC 5165
QY 4991 AGAGAGGAGAGCGGAGAGAGAGATGTCTGAAATCTTACCTGACCGGACTCTGAGCA 5050
Db 5166 GCGAGGAGAGAGCTGAGAGAGATGTCTCGAGATCTTACCTGACAGCGGCTGTGGCA 5225
QY 5051 CTAAAGGCACTGCAAGAGTTTGTGATGAGCTCTTTGAGACCACTTCAAGAGCGGCA 5110

Db	5226	CCAAAGGCAACCTGCAGAGATTCGTGATGAGACTCTTTGAGACAGTGTTCAGCAAGGCC	5285
Qy	5111	ACCGTGGCTCTGCCCTGCCCCCTGGCCATCAAGTACATGTTTGACTCTCTGGATGAGCAGG	5170
Db	5286	ACCAGGGCTCGGCCCTGCCCTCGGCCCAACAGTACATGTTGACTCTCTGGATGAGCAGG	5345
Qy	5171	CTGATTAACATGAGCATTCATGACCCGCAGTCCGCCATACCTCGGAGAGACAAATTGCTGC	5230
Db	5346	CGGACAGGCGCAGATCAGGAGACCCCGATGTGGCGCACACTGGAAAGGACAACTGCTGC	5405
Qy	5221	CCCTGAGGTTTTGGGTCAACATGATCAAGAACCCGCACTTTGTGTTGACATCCATAGA	5290
Db	5406	CGCTGCGCTTTGGGTGATGTGATCAAGAACCCGCACTTTGTGTTGACATCCACAGA	5465
Qy	5291	ACAGCATCACAGACGCGCTCTCTGTGGTGGGCTCAGACCTTCATGAGACTCTTGCTCA	5350
Db	5466	ACAGCATCACGGAATGCCCTGTGCTGTGGTGTGAGCCAGACCTTCATGAGACTCTTGCTCA	5525
Qy	5351	CGTCAGAGCACCGGCTGGGCAAGGACTCGCCCTCCACACAGCTGCTGTATGCCAAGACA	5410
Db	5526	CATCCGAGCACCGGCTGGGGAAGGACTCGCCCTCCAAACACTGTCATCCCAAGACA	5585
Qy	5411	TCCCCAGGTCACAGATTTGGGTGGAGAGATTTTACTCAGACATGAGGGAAGATGCGACCA	5470
Db	5586	TCCCCACATCAAGAGCTGGGTGGAGAGATTTATTCGAGACATTTGCAGAAATGGCATCCA	5645
Qy	5471	TCAGGACCAAGACATGAAACGATACCTGSGCTGAGACAGTCCCGGATGACATGAAATGCT	5530
Db	5646	TCAGGACCAAGACATGATGCTTACCTGSGTGGAGCACTCCCGCTCCACGCCACAGCGCT	5705
Qy	5531	TCAACACCATGAGTGCACCTCTCAGAGATCTTCTCTATGTGGGCAAAATACAGCAGAGCA	5590
Db	5706	TCAGGCTCTGATGTGCGCTCAACGAGCTGTATTTTATAGTCACCAAGTACCGCCAGAGCA	5765
Qy	5591	TCCTTGAACCTCTGGACCAAGATGACCAAGTGTGGAGACAGAACTGGCTTCAAACTAG	5650
Db	5766	TTCTTCACGGCTCTGGACCGAGATGCTCTTTGTGCGAACAATTAAGTTGGCGAGAACTGG	5825
Qy	5651	AACCAAGTCATAACCTCANTGAGCTTAAACAGCTGA	5685
Db	5826	AACCAATATCAGCTTCTGTTCACGACAGCTTAA	5860

RESULT	8	
ABV76746		
ID	ABV76746	standard; cDNA; 3914 BP.
XX		
AC	ABV76746;	
XX		
DT	28-MAR-2003	(first entry)
XX		
DE	Human plexin 2 protein 56.11-encoding cDNA.	
XX		
KM	Human; plexin 2 protein 56.11; recombinant production; gene therapy;	
XX	palsy; dementia; gene; ss.	
OS	Homo sapiens.	
XX		
FT	Key	Location/Qualifiers
FT	1391..2926	
CDS	/*tag= a	
FT	/product= "Human plexin 2 protein 56.11"	
XX		
EN	CN1359955-A.	
XX		
PD	24-JUL-2002.	
XX		
PF	20-DEC-2000; 2000CN-00135177.	
XX		
PR	20-DEC-2000; 2000CN-00135177.	
XX		
PA	(BODE-) BODE GENE DEV CO LTD SHANGHAI.	

Pt	Mao Y., Xie Y;
Xx	WPJ, 2002-73643/80.
Dn	P-PDS5; ABP58529.
Pt	Polypeptide-human plexin 2 protein 56.11 and polynucleotide for coding it.
Pt	Claim 6; Page 24-26 (Disclosure); 35pp; Chinese.
Xx	The invention relates to human plexin 2 protein 56.11 (ABP58529) and nucleic acids encoding it (ABV76746). The protein has a molecular weight of 56.11 kD. The invention also relates to a method for the recombinant production of the protein, an antagonist of the protein, and the use of the protein, gene and antagonist in therapeutic applications. Plexin 2 protein 56.11 can be used in the treatment of a variety of diseases such as palsy and dementia. The present sequence represents cDNA encoding human plexin 2 protein 56.11
Sq	Sequence 3914 BF; 924 A; 1181 C; 1016 G; 793 T; 0 U; 0 Other;
Qy	Query Match 17.1%; Score 974; DB 6; Length 3914; Best Local Similarity 67.4%; Pred. No. 2.2e-224; Matches 1527; Conservative 0; Mismatches 510; Indels 227; Gaps 2;
Db	3023 CCACATCCTCAGATGAGGTGCTAAGAATGAAGTGTGGTGTCAGGTGAGCACAGGCCAAAGA 3082 CCCCATCATCAATTGGCGTTGGCCGGTCCCTGTTCTGTGAGAGTGCAGCCGACCATG 1475
Oy	3083 TCACACCGAGACTTGGTCTTTCAGTAATGTGAAGAACCCACCATCTGCGGGATTGAGCCAG 3142
Dd	1476 TGAGTAGCAACTGTCAGATTGAGTACAATAATAGACCTCGGGTTCACAGCGATGAGCCAG 1535
Oy	3143 AATGAGCACTTTGTACGTGAGAAAACACCCATCGCCGTATGGGGAGACCCACCTGAGACTCA 3202
Dd	1536 AGTGGACACTTGGCCAGTGGCCAACCCCCTGACATCAAGGCTTCAACTGGAATGTCA 1595
Oy	3203 TAACGAAACCCCGAGATCCGTGGCCAGATGAGAGAGGAAAGAGACATTAATCTGTAGG 3262
Dd	1596 TTCAGGAGCCAGAGATCCAGATCAAATTAATGACAAAGATCTGTCAATGTGTGTAAG 1655
Oy	3263 TTTCGAAACGCTACGTGATGATGACCTGTGAGGCGCCGCGCTCGCTGGGTCTGTAGCCACC 3322
Dd	1656 TTGTGAACAACAACCACTTCACTCTGTGGCACTCTGTGACCAAGCATCAOCCGCTG 1715
Oy	3323 AGTCAGACCTGAGCCGAGAGGCGCCGAGAGTTTGGCTTCACTCTGAGACAAGTCCAGTCC 3382
Dd	1716 GCCTGGACACTGTGTGAAGCGCCACAGATAGTTGTGATTGTCTTMAAAGTCCAAATCCT 1775
Oy	3383 TGCTCAATCCCAACAAGAACAACCTGCACCTAATCCCAACCGGGTGTGAGGCGCTTTG 3442
Dd	1776 TGTCAATTTTACAAGACAACCAAGTTTATCTAACHACCCAAACCCGACTTTGAACTGTTA 1835
Oy	3443 GTCCCTCAGAAATCCTGAGGTCAAGCCTGAGCAGCCCATCATCTTAAAGGGCAAGAAC 3502
Dd	1836 GCCCTATCGAGAGTCTTGATCAAAAGCAGAGATGCCCATCATCTGTAAGGGCAAAAACC 1895
Oy	3503 TGAATCCGCGCTGTGTGGCTGGGGGCAAGTGAAGCTGAACCTACATCTGTTGGTTGGGGAGA 3562
Dd	1896 TCTGCGCTCTCTGCGCTCTGGAGG---GGCCAAACTCAACTACATCTGCTCACTGGAAGA 1952
Oy	3563 AGCGGTGACCGGTACCGGTGTGACATATCAAGTGTCTCTGAGAGTCCGCCCAACTCTACCG 3622
Dd	1953 CCCCTTGTGTGTACCCGTAATGAGAACCAAGCTTCTCTGGAGGCTTCCCAACTCAACCG 2012
Oy	3623 GAAGGCAAAAGTATGGCCCCGTGTGGTGCATGAGTACTCCCCGGGGATGTGTACA 3682
Dd	2013 GGAGACACAAAGTATGATGTTCACTGTGGCGGGAAGTGTGTTCTGCTGGCTGTGGTGAAGTG 2072
Oy	3683 TTGGCCCCGGAAGCGCGCTGAGCTGTGCCCGGCAATCTGCACATGTGACGTGGCTTGCGCGCC 3742
Dd	2073 TCAATTCAGACAGTTTGTACTACCTGTGCAAGCATGTCTACAGATGCGGCGCGCGGCAAGCC 2132

QY 3743 TCTCATCATTTTTCATGCTGGCCGCTGCTCATTCGCTTAAAGCAAGTCCCGGAAAGTG 3802
 Db 2113 TCTCTCATCATGTCATGCTGCTCATTCGCTTAAAGCAAGTCCCGGAAAGTG 2192
 QY 3803 ACCTCAGCTGAGGCTGAGATGAGATGAGCAACCTGAGTCCCGTGTGGCCCTGG 3862
 Db 2193 ACCTCAGCTGAGGCTGAGATGAGATGAGCAACCTGAGTCCCGTGTGGCCCTGG 2252
 QY 3863 AGTGAAGAAAGCTTTTCCGAGCTGAGAGAGAGATCATCATGAGTGAACAGTCTGG 3922
 Db 2253 AGTGAAGAAAGCTTTTCCGAGCTGAGAGAGAGATCATCATGAGTGAACAGTCTGG 2312
 QY 3923 ATGAGAGGAGATTCGCTTCCGAGCTGAGATGAGATGAGTGAACAGTCTGG 3982
 Db 2313 ACCGCTCAGAAATCCCTTAACTGAGATGAGTGAACAGTCTGG 2372
 QY 3983 GAATTGAAGACCACTGCTGCTGAGAGCTTGAAGTCCCGGCTACCGGAGAGCTGG 4042
 Db 2373 GCATGAGAGACCACTGCTGCTGAGAGCTTGAAGTCCCGGCTACCGGAGAGCTGG 2432
 QY 4043 TGGAAAGAGCTGAGATGCTTCCGAGCTGAGATGAGATGAGTGAACAGTCTGG 4102
 Db 2433 TGGAAAGAGCTGAGATGCTTCCGAGCTGAGATGAGATGAGTGAACAGTCTGG 2492
 QY 4103 TCATCCGAGCTGAGATGCTTCCGAGCTGAGATGAGATGAGTGAACAGTCTGG 4162
 Db 2493 TCATCCGAGCTGAGATGCTTCCGAGCTGAGATGAGATGAGTGAACAGTCTGG 2552
 QY 4163 CACTCATATGAGCTGCTGAGAGAGAGATGAGATGAGTGAACAGTCTGG 4222
 Db 2553 CGCTCATATGAGCTGCTGAGAGAGAGATGAGATGAGTGAACAGTCTGG 2612
 QY 4223 TGTGAGAGCTGCTGAGAGAGAGATGAGATGAGTGAACAGTCTGG 4282
 Db 2613 TGTGAGAGCTGCTGAGAGAGAGATGAGATGAGTGAACAGTCTGG 2672
 QY 4283 GAGAGAGCTGAGATGCTGAGAGAGAGATGAGATGAGTGAACAGTCTGG 4342
 Db 2673 GAGAGAGCTGAGATGCTGAGAGAGAGATGAGATGAGTGAACAGTCTGG 2732
 QY 4343 AGTTCCTCA----- 4351
 Db 2733 AGTTCCTCAAGTGAAGAGAGAGATGAGATGAGTGAACAGTCTGG 2792
 QY 4351----- 4351
 Db 2793 GCTGCTTGAATCTGAGAGAGAGATGAGATGAGTGAACAGTCTGG 2852
 QY 4352----- 4352
 Db 2853 GAGAGCTGCTGAGAGAGAGATGAGATGAGTGAACAGTCTGG 2912
 QY 4352-----AGAGTG 4358
 Db 2913 AGAGTGAAGAGAGAGAGATGAGATGAGTGAACAGTCTGG 2972
 QY 4359 TGTGAGAGAGAGAGAGATGAGATGAGTGAACAGTCTGG 4418
 Db 2973 TGTGAGAGAGAGAGAGATGAGATGAGTGAACAGTCTGG 3032
 QY 4419 CATTTGAGAGAGAGAGAGATGAGATGAGTGAACAGTCTGG 4478
 Db 3033 CATTTGAGAGAGAGAGAGATGAGATGAGTGAACAGTCTGG 3092
 QY 4479 GAGAGATGAGATGAGATGAGATGAGTGAACAGTCTGG 4538
 Db 3093 GAGAGATGAGATGAGATGAGATGAGTGAACAGTCTGG 3152
 QY 4539 CGAGATGAGATGAGATGAGATGAGTGAACAGTCTGG 4598
 Db 3153 AGAGATGAGATGAGATGAGATGAGTGAACAGTCTGG 3212

QY 4599 GGATGCACTTCAAGAAATGCTTGTCTCCACCGGCGCAAGCTGAGATGATGATCT 4658
 Db 3213 TGAATGCTGATTAAGAAATGCTTGTCTCCACCGGCGCAAGCTGAGATGATGATCT 3272
 QY 4659 GAGATGCGCAAGAAAGTGGGGCAAGATGATCTTGAAGATGAAGATCAACCA 4718
 Db 3273 GAGATGCGCGCAAGGCGGAGATGCGGGGATGCTGAGATGAGATCAACCA 3332
 QY 4719 GATTGAAATGATTGAAGAGCACTGAACACACTGGCCCACTAACAGTGGATGTTT 4778
 Db 3333 GATTGAGGATGATGAGAGAGGCTCAACACTGATGATTAATACAGTGTGAGACAGGTC 3392
 QY 4779 CGTGTGGATTAATGATTCAGAGAGATGAGATGATGATGATGATGATGAT 4838
 Db 3393 GGTGTGGCTGTGTCTCCCAAGAGATGATGATGATGATGATGATGATGAT 3452
 QY 4839 CTCCAGAGCTGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 4898
 Db 3453 CTCCAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3512
 QY 4899 CCTCCGCTCAGAGACACTGATGATGATGATGATGATGATGATGATGATGAT 4958
 Db 3513 CCTCCGCTCAGAGACACTGATGATGATGATGATGATGATGATGATGATGAT 3572
 QY 4959 CCTAGTGAAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5018
 Db 3573 TCTGTGAAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3632
 QY 5019 GTCTGAATCTACTGATGATGATGATGATGATGATGATGATGATGATGAT 5062
 Db 3633 GTCTGAATCTACTGATGATGATGATGATGATGATGATGATGATGATGAT 3676

RESULT 9
 ABK70006
 ID ABK70006 standard; DNA; 2597 BP.
 AC ABK70006;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE cDNA encoding human Pro peptide #46.
 XX
 KM Human; ss; gene; PRO; secreted protein; transmembrane protein;
 KM genetic disorder; tumor; cancer.
 OS Homo sapiens.
 XX
 PD WO20022488-A2.
 XX
 PD 28-MAR-2002.
 XX
 PF 29-AUG-2001; 2001MO-US027099.
 XX
 ER 01-SEP-2000; 2000US-0229896P.
 ER 05-SEP-2000; 2000US-0230621P.
 ER 22-SEP-2000; 2000US-0235147P.
 ER 10-NOV-2000; 2000MO-US030873.
 ER 12-JAN-2001; 2001US-0261878P.
 ER 16-JAN-2001; 2001US-0261910P.
 ER 16-JUN-2001; 2001US-0261939P.
 ER 16-JUN-2001; 2001US-0262150P.
 ER 25-JAN-2001; 2001US-0264395P.
 ER 02-FEB-2001; 2001US-0264421P.
 ER 09-FEB-2001; 2001US-0267623P.
 ER 28-FEB-2001; 2001MO-US006520.
 ER 09-MAR-2001; 2001US-0274399P.
 ER 03-APR-2001; 2001US-0280982P.
 ER 04-APR-2001; 2001US-0282129P.
 ER 04-APR-2001; 2001US-0282199P.
 ER 09-MAY-2001; 2001US-0290899P.
 ER 25-MAY-2001; 2001MO-US017092.
 ER 01-JUN-2001; 2001MO-US017800.

PR 20-JUN-2001; 2001MO-US019692.
 PR 29-JUN-2001; 2001MO-US021066.
 PR 09-JUL-2001; 2001MO-US021735.
 XX
 (GETH) GENENTECH INC.
 PI Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC,
 PI Gurney AL, Smith V, Stephan J, Watanabe CK, Wood WI, Zhang Z,
 PI Fong S;
 XX WPI: 2002-362426/39.
 DR P-PSDB; ABG34075.
 XX
 PT New PRO polypeptides and polynucleotides encoding the polypeptides,
 PT useful in gene therapy, chromosome identification, tissue typing, or for
 PT genetic analysis of individuals with genetic disorders.
 XX
 PS Claim 2; Fig 91; 218pp; English.
 CC This invention relates to the cDNA and protein sequences of novel
 CC secreted and transmembrane polypeptides PRO polypeptides. The invention
 CC also comprises a method for producing the proteins of the invention by
 CC recombinant means and antibodies specific for the protein of the
 CC invention. The antibody may be used for detecting the PRO proteins of the
 CC invention and may be used to modify their activity. polynucleotides may
 CC be used as hybridisation probes for a cDNA library to isolate the full-
 CC length PRO cDNA or to isolate other cDNAs, to construct hybridisation
 CC probes for mapping the gene which encodes that PRO and for genetic
 CC analysis of individuals with genetic disorders, in assays to identify
 CC other proteins or molecules involved in binding reaction, to generate
 CC transgenic animals or knock-out animals which in turn are useful in the
 CC development and screening of therapeutically useful reagents, for
 CC chromosome identification, and tissue typing. The PRO polypeptides are
 CC useful in gene therapy, and as molecular weight markers for protein
 CC electrophoresis purposes. The sequences may also be used to detect
 CC overexpression on PRO polypeptides in cancerous tumours and for screening
 CC for differentially expressed genes using microarray technology. The
 CC present sequence represents a cDNA encoding a human PRO protein of the
 CC invention
 CC
 XX
 SQ Sequence 2597 BP; 688 A; 570 C; 642 G; 697 T; 0 U; 0 Other;
 Query Match 13.9%; Score 793.8; DB 6; Length 2597;
 Best Local Similarity 99.7%; Pred. No. 6; 1e-181;
 Matches 799; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 579 CGAGTATTTTCCCACTCTCCAGCCGGAACCTGACCAAGAACTTGGAGCGGATGGCAT 638
 DB 1 CGAGTATTTTCCCACTCTCCAGCCGGAACCTGACCAAGAACTTGGAGCGGATGGCAT 60
 QY 639 GTTCGCGTACGTCCTTCATGATGATGTCGTGCGCTCGATGATTAAGTCCCTTCGACAC 698
 DB 61 GTTCGCGTACGTCCTTCATGATGATGTCGTGCGCTCGATGATTAAGTCCCTTCGACAC 120
 QY 699 CTTCACCATCATCCCTGACCTTGTATCTATCTATGTCATGCTTTTGAAGTGGCAACTT 758
 DB 121 CTTCACCATCATCCCTGACCTTGTATCTATCTATGTCATGCTTTTGAAGTGGCAACTT 180
 QY 759 TGTCTACTTTTGAACCTTCACCTGAGATGATGTCCTCCACAGGCTCCACCAAGGA 818
 DB 181 TGTCTACTTTTGAACCTTCACCTGAGATGATGTCCTCCACAGGCTCCACCAAGGA 240
 QY 819 GCAAGGTATATACATCAAGCTCGTGAAGCTTTTGAAGAGAGACACAGCTTCAACTCTTA 878
 DB 241 GCAAGGTATATACATCAAGCTCGTGAAGCTTTTGAAGAGAGACACAGCTTCAACTCTTA 300
 QY 879 TGTAGAGGTGCGCATTTGGCTGTAGAGCGAGTGGGCTGAATACCGCTGTGACGCTGC 938
 DB 301 TGTAGAGGTGCGCATTTGGCTGTAGAGCGAGTGGGCTGAATACCGCTGTGACGCTGC 360
 QY 939 CTACCTGTCCAAACGGGGGGCGGTGCTTGGCAGGACCTTTGGAGTCCATCCAGATATGA 998
 DB 361 CTACCTGTCCAAACGGGGGGCGGTGCTTGGCAGGACCTTTGGAGTCCATCCAGATATGA 420

QY 999 CTTGCTCTTCAACCGTCTTCTTCCAAAGGCGCAAGAGCGGAAATGAAATCCTGATGATGTC 1058
 DB 421 CTTGCTCTTCAACCGTCTTCTTCCAAAGGCGCAAGAGCGGAAATGAAATCCTGATGATGTC 480
 QY 1059 GGGCCCTGATGATCTTCACTTCTTGAAGCAGATTAATGACCGGATTAAGGAGCGGCTGACGTC 1118
 DB 481 GGGCCCTGATGATCTTCACTTCTTGAAGCAGATTAATGACCGGATTAAGGAGCGGCTGACGTC 540
 QY 1119 TTGTTACCGGGGCGAGGGCAGCGCTGACCTGCGCTTCAAGGTGAAGGACATCCCTCTG 1178
 DB 541 TTGTTACCGGGGCGAGGGCAGCGCTGACCTGCGCTTCAAGGTGAAGGACATCCCTCTG 600
 QY 1179 CAGCAGTGGCGCTTCAACATGATGACGATTAATCTGCTGAGCTGACATGATGCTCCCT 1238
 DB 601 CAGCAGTGGCGCTTCAACATGATGACGATTAATCTGCTGAGCTGACATGATGCTCCCT 660
 QY 1239 GGGAGTGTCCGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1298
 DB 661 GGGAGTGTCCGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
 QY 1299 GTCTGTATGCGCATATGCTTCAAGAACCACTCTGCTGCTTGTGGGCAACCAAGTGG 1358
 DB 721 GTCTGTATGCGCATATGCTTCAAGAACCACTCTGCTGCTTGTGGGCAACCAAGTGG 780
 QY 1359 CAAGCTGAGAAGATTC 1375
 DB 781 CAAGCTGAGAAGATTC 797
 RESULT 10
 ID ADA01359
 ADADA01359 standard; cDNA; 2597 BP.
 XX
 XX ADA01359;
 XX
 XX 06-NOV-2003 (first entry)
 XX
 XX
 DE Human PRO polynucleotide #46.
 DE
 XX
 KW Human; gene; ss; PRO; secreted polypeptide; transmembrane polypeptide;
 KW tumor necrosis factor-alpha; TNF-alpha; blood; chondrocyte cell; tumour;
 KW adrenal; lung; colon; breast; prostate; rectum; cervix; liver; cancer;
 KW microvascular endothelial cell; endothelial cell tube formation;
 KW sports-related joint problem; articular cartilage defect; osteoarthritis;
 KW rheumatoid arthritis; osteopathic; antirheumatic; antirheumatic.
 KW
 OS Homo sapiens.
 XX
 XX US2003068779-A1.
 XX
 XX 10-APR-2003.
 XX
 XX 16-SEP-2002; 2002US-00245107.
 XX
 XX 09-MAY-2001; 2001US-0290589P.
 XX 29-AUG-2001; 2001MO-US027099.
 XX 18-JUL-2002; 2002US-00197942.
 XX
 XX (GETH) GENENTECH INC.
 XX
 XX Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC,
 XX Gurney AL, Smith V, Stephan J, Watanabe CK, Wood WI, Zhang Z,
 XX Fong S;
 XX WPI: 2003-625484/59.
 XX P-PSDB; ADA01360.
 XX
 XX Novel isolated PRO1313, PRO20080 or PRO21383 polypeptide useful for
 XX stimulating proliferation of human microvascular endothelial cells, and
 XX PRO6018 polypeptide useful for stimulating proliferation of chondrocyte
 XX cells.

Claim 2; Fig 91; 307pp; English.

The invention relates to isolated human PRO polypeptides (secreted and transmembrane polypeptides) and the polynucleotides encoding them. The invention also relates to an antibody which specifically binds to a PRO polypeptide, a method for stimulating the release of tumour necrosis factor-alpha (TNF-alpha) from human blood, a method for stimulating the proliferation or differentiation of chondrocyte cells and a method for detecting the presence of a tumour in a mammal (e.g. adrenal, lung, colon, breast, prostate, rectal, cervical and liver tumours). The polynucleotides are useful in molecular biology, including uses as hybridisation probes, in chromosome and gene mapping, in generating antisense RNA and DNA and in gene therapy. The polynucleotides may also be used in preparing PRO polypeptides by recombinant techniques and in generating either transgenic animals or knock-out animals which are useful in the development and screening of therapeutically useful reagents. The PRO polypeptides or antibodies are used in preparing a medicament for treating a condition responsive to the polypeptides or antibodies, such as tumours, for stimulating and inhibiting proliferation of human microvascular endothelial cells, for inducing endothelial cell tube formation and for treating sports-related joint problems, articular cartilage defects, osteoarthritis and rheumatoid arthritis. This sequence represents a human PRO polynucleotide of the invention.

Sequence 2597 BP; 688 A; 570 C; 642 G; 697 T; 0 U; 0 Other;

Query Match 13.9%; Score 793.8; DB 8; Length 2597; Best Local Similarity 99.7%; Pred. No. 6,1e-181; Matches 795; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

579 CGAGTATTTTCCACCATCTCCAGCCGGAACCTGAGCGCGATGGCAT 638
1 CGAGTATTTTCCACCATCTCCAGCCGGAACCTGAGCGCGATGGCAT 60
639 GTTCGGGTACGCTTTCATGATGAGTTCGTGGCTCGATGATTAAGATCCCTCGGACAC 698
61 GTTCGGGTACGCTTTCATGATGAGTTCGTGGCTCGATGATTAAGATCCCTCGGACAC 120
699 GTTCACCATCATCCTGACCTTGAATGATGATGATGATGATGATGATGATGATGATGAT 758
121 GTTCACCATCATCCTGACCTTGAATGATGATGATGATGATGATGATGATGATGATGAT 180
759 GTTCACCATCATCCTGACCTTGAATGATGATGATGATGATGATGATGATGATGATGAT 818
181 GTTCACCATCATCCTGACCTTGAATGATGATGATGATGATGATGATGATGATGATGAT 240
819 GCGAGTGTATACATCCAGCTCGTGAAGGCTTTGCAAGAGAGACAGAGCTTCACTCTCTA 878
241 GCGAGTGTATACATCCAGCTCGTGAAGGCTTTGCAAGAGAGAGAGAGCTTCACTCTCTA 300
879 TGTAGAGGTGCGCATGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 938
301 TGTAGAGGTGCGCATGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
939 CTACCTGTGTCCAAAGCGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 998
361 CTACCTGTGTCCAAAGCGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
999 CCGGCTGTGTCCAAAGCGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1058
421 CCGGCTGTGTCCAAAGCGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
1059 GCGCTGTGTCCAAAGCGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1118
481 GCGCTGTGTCCAAAGCGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
1119 TTTTGTACCGGGGAG 1178
541 TTTTGTACCGGGGAG 600
1179 CAGCAGGTGCGCTCTTAAACATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 1238
601 CAGCAGGTGCGCTCTTAAACATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 660

1239 GGGAGTGTCCGAGATGCTGCTGAGATTTCCCGTTTACAGAGAGACAGGAGCCGATGAC 1298
661 GGGAGTGTCCGAGATGCTGCTGAGATTTCCCGTTTACAGAGAGAGAGAGCCGATGAC 720
1299 GTCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1358
721 GTCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
1359 CAGCTGTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1375
781 CAGCTGTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 797

RESULT 11

ID ADA43788 standard; cDNA; 2597 BP.
XX ADA43788;
AC ADA43788;
XX 20-NOV-2003 (first entry)
XX Human cDNA encoding secreted/transmembrane polypeptide PRO34003.

XX ss; gene; human; PRO; secreted protein; transmembrane protein;
XX endothelial cell tube formation; chondrocyte cell differentiation;
XX microvascular endothelial cell; tumour; lung tumour; colon tumour;
XX breast tumour; prostate tumour; rectal tumour; kidney tumour;
XX liver tumour; cytostatic; vaccine.

XX Homo sapiens.

XX US2003064474-A1.

XX 03-APR-2003.

XX 16-SEP-2002; 2002US-00245859.

XX 29-AUG-2001; 2001WO-US027059.

XX 18-JUL-2002; 2002US-00197942.

XX (GENT) GENENTECH INC.

XX Baker KP, Baton DV, Filvaroff B, Goddard A, Grimaldi JC;
XX Gurney AL, Smith V, Stephan JP, Watanabe CK, Wood WI, Zhang Z;
XX Fong S;

XX WPI; 2003-605867/57.

XX P-PSDB; ADA43789.

XX New isolated nucleic acid encoding a PRO polypeptide, e.g. PRO20080 or
XX PRO2183, useful in molecular biology, chromosome and gene mapping, in
XX generating antisense RNA and DNA, and in gene therapy.

XX Claim 2; Fig 91; 308pp; English.

XX The invention relates to an isolated secreted/transmembrane (PRO)
XX polypeptide, having at least 80% sequence identity to a sequence selected
XX from any one of the 57 amino acid sequences given in specification, or to
XX a sequence encoded by a nucleic acid molecule selected from any one of
XX the nucleic acids deposited under any of the ATCC accession numbers given
XX in specification, or a sequence having at least 80% identity to PRO
XX lacking its associated signal peptide, an extracellular domain of PRO
XX with or without its associated signal peptide. Also included are vectors,
XX transformed host cells, anti-PRO antibodies, the nucleic acids encoding
XX PRO, PRO fusion proteins, inducing endothelial cell tube formation (by
XX administering PRO281, PRO1560, PRO189, PRO4499, PRO6308, PRO6000,
XX PRO10275, PRO21207, PRO20933 or PRO34274 polypeptide or its agonist) and
XX an oligonucleotide probe derived from any one of the above nucleotide
XX sequences. PRO6018 polypeptide is useful for stimulating the
XX proliferation or differentiation of chondrocyte cells. PRO1317, PRO20080
XX and PRO2183 polypeptides are useful for stimulating the proliferation of
XX human microvascular endothelial cells. PRO6071, PRO4487 and PRO6006

CC polypeptides are useful for inhibiting the proliferation of human
CC microvascular endothelial cells. PRO polypeptides are useful for
CC detecting the presence of tumour in a mammal, including tumours of lung,
CC colon, breast, prostate, rectal, kidney and liver. PRO281, PRO1560,
CC PRO189, PRO499, PRO6308, PRO6000, PRO10275, PRO21207, PRO20933 and
CC PRO34274 polypeptides are useful for inducing endothelial cell tube
CC formation. PRO or the antibody are useful in the preparation of a
CC medication for treating a condition responsive to PRO polypeptide. The
CC oligonucleotide probes are useful for isolating genomic and cDNA
CC nucleotide sequences, for measuring or detecting the expression of an
CC associated gene, and as antisense probes. PRO nucleic acid is useful as a
CC hybridisation probe, in chromosome and gene mapping, in the generation of
CC antisense RNA and DNA, and for the preparation PRO polypeptides. PRO and
CC PRO nucleic acid are useful as therapeutic agents, e.g. vaccines. The
CC present sequence encodes a PRO protein.

XX Sequence 2597 BP; 688 A; 570 C; 642 G; 697 T; 0 U; 0 Other;

XX Query Match 13.9%; Score 793.8; DB 8; Length 2597;

XX Best Local Similarity 99.7%; Pred. No. 6.1e-181;
XX Matches 795; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 579 CGAGTATTTTCCACCATCTCCAGCCGGAGAACTGACAGAACTCTGAGCGGATGCGAT 638
DB 1 CGAGTATTTTCCACCATCTCCAGCCGGAGAACTGACAGAACTCTGAGCGGATGCGAT 60
QY 639 GTTCGCGTACGCTCTTCATGATAGTTCGTCGCTCGATGATTAAGATCCCTTCGACAC 698
DB 61 GTTCGCGTACGCTCTTCATGATAGTTCGTCGCTCGATGATTAAGATCCCTTCGACAC 120
QY 699 CTTACACCATATCTCTGATCTTATATCTAATCTAATCTAATCTAATCTAATCTAATCT 758
DB 121 CTTACACCATATCTCTGATCTTATATCTAATCTAATCTAATCTAATCTAATCTAATCT 180
QY 759 TGTCTACTTTTACCCCTCCACCTGAGATGCTGTCTCAGACAGGCTCCACACACAGGA 818
DB 181 TGTCTACTTTTACCCCTCCACCTGAGATGCTGTCTCAGACAGGCTCCACACACAGGA 240
QY 819 GCAGGTGTATATCCAAAGCTGTGAGGCTTTGCAAGAGACAGCCTTCACTCTCA 878
DB 241 GCAGGTGTATATCCAAAGCTGTGAGGCTTTGCAAGAGACAGCCTTCACTCTCA 300
QY 879 TGTAGAGGTGCGCATTCGCTGAGGCGATGAGGCGATGAGGCGATGAGGCGATGAGG 938
DB 301 TGTAGAGGTGCGCATTCGCTGAGGCGATGAGGCGATGAGGCGATGAGGCGATGAGG 360
QY 939 CTACCTGTCCAAAGCGGGGCGCTGTGACAGACCTTTGAGTCCATCCAGATGATGA 998
DB 361 CTACCTGTCCAAAGCGGGGCGCTGTGACAGACCTTTGAGTCCATCCAGATGATGA 420
QY 999 CTTGCTCTTACCGCTCTTCTCCAAAGGCGGAGCGGAAATTAATCCCTGGATGAGTC 1058
DB 421 CTTGCTCTTACCGCTCTTCTCCAAAGGCGGAGCGGAAATTAATCCCTGGATGAGTC 480
QY 1059 GGCCTGTGATCTTCACTTTGAAGCAGATTAATGACCGCATTAAGAGCGGCTGACATC 1118
DB 481 GGCCTGTGATCTTCACTTTGAAGCAGATTAATGACCGCATTAAGAGCGGCTGACATC 540
QY 1119 TTGTATCCGGGGGAGAGGCGCTGACCTGCTGCTGCTTAAGGTGAAGACATCCCTCG 1178
DB 541 TTGTATCCGGGGGAGAGGCGCTGACCTGCTGCTGCTTAAGGTGAAGACATCCCTCG 600
QY 1179 CAGCAGTGCCTCTTAACATTAACATTAACATTAACATTAACATTAACATTAACATTA 1238
DB 601 CAGCAGTGCCTCTTAACATTAACATTAACATTAACATTAACATTAACATTAACATTA 660
QY 1239 GGGAGTGTCCGACATGTGTGCTGAATTCCTGCTTCAAGGAGACAGGAGCCGATGAC 1298
DB 661 GGGAGTGTCCGACATGTGTGCTGAATTCCTGCTTCAAGGAGACAGGAGCCGATGAC 720
QY 1299 GTCTGTATGCAATATGTCTACAAAGACACTCTGTGGCTTTGTGGGACCAAAAGTGG 1358
DB 721 GTCTGTATGCAATATGTCTACAAAGACACTCTGTGGCTTTGTGGGACCAAAAGTGG 780

QY 1359 CAAAGTGAAGAAGATCC 1375
DB 781 CAAAGTGAAGAAGTGC 797

RESULT 12

ADA43556
ID ADA43556 standard; cDNA; 2597 BP.

AC ADA43556;

DT 20-NOV-2003 (first entry)

DE Human cDNA encoding secreted/transmembrane polypeptide PRO34003.

XX ss; gene; human; PRO; secreted protein; transmembrane protein;
XX endothelial cell tube formation; chondrocyte cell differentiation;
XX microvascular endothelial cell; tumour; lung tumour; colon tumour;
XX breast tumour; prostate tumour; rectal tumour; kidney tumour;
XX liver tumour; cytostatic; vaccine.

OS Homo sapiens.

PN US2003073196-A1.

PD 17-APR-2003.

PF 18-SEP-2002; 2002US-00246210.

PR 04-APR-2001; 2001US-0282199P.

PR 29-AUG-2001; 2001WO-US027099.

FR 18-JUL-2002; 2002US-00197942.

XX (GENTH) GENENTECH INC.

XX Baker KP, Eaton DV, Filvaroff E, Goddard A, Grimaldi JC,
XX Gurney AL, Smith V, Stephan JP, Watanabe CK, Wood WI, Zhang Z,
XX Fong S;

XX WPI; 2003-743814/70.

XX P-P5DB; ADA43557.

XX New secreted and transmembrane PRO polypeptides, e.g. PRO20080 or
XX PRO31383 useful for stimulating the proliferation or differentiation of
XX chondrocyte cells and detecting the presence of a tumor in a mammal.

XX Claim 2; Fig 91; 307pp; English.

XX The invention relates to an isolated secreted/transmembrane (PRO)
XX polypeptide, having at least 80% sequence identity to a sequence selected
XX from any one of the 57 amino acid sequences given in specification, or to
XX a sequence encoded by a nucleic acid molecule selected from any one of
XX the nucleic acids deposited under any of the ATCC accession numbers given
XX in specification, or a sequence having at least 80% identity to PRO
XX lacking its associated signal peptide, an extracellular domain of PRO
XX with or without its associated signal peptide. Also included are vectors,
XX transformed host cells, anti-PRO antibodies, the nucleic acids encoding
XX PRO, PRO fusion proteins, inducing endothelial cell tube formation (by
XX administering PRO281, PRO1560, PRO189, PRO499, PRO6308, PRO6000,
XX PRO10275, PRO21207, PRO20933 or PRO34274 polypeptide or its agonist) and
XX an oligonucleotide probe derived from any one of the above nucleotide
XX sequences. PRO6018 polypeptide is useful for stimulating the
XX proliferation or differentiation of chondrocyte cells. PRO313, PRO20080
XX and PRO31383 polypeptides are useful for stimulating the proliferation of
XX human microvascular endothelial cells. PRO6071, PRO4487 and PRO6006
XX polypeptides are useful for inhibiting the proliferation of human
XX microvascular endothelial cells. PRO polypeptides are useful for
XX detecting the presence of tumour in a mammal, including tumours of lung,
XX colon, breast, prostate, rectal, kidney and liver. PRO281, PRO1560,
XX PRO189, PRO499, PRO6308, PRO6000, PRO10275, PRO21207, PRO20933 and
XX PRO34274 polypeptides are useful for inducing endothelial cell tube
XX formation. PRO or the antibody are useful in the preparation of a

Medicament for treating a condition responsive to PRO polypeptide. The CC oligonucleotide probes are useful for isolating genomic and cDNA CC nucleotide sequences, for measuring or detecting the expression of an CC associated gene, and as antisense probes. PRO nucleic acid is useful as a CC hybridisation probe, in chromosome and gene mapping, in the generation of CC antisense RNA and DNA, and for the preparation PRO polypeptides. PRO and CC PRO nucleic acid are useful as therapeutic agents, e.g. vaccines. The CC present sequence encodes a PRO protein.

XX Sequence 2597 BP; 688 A; 570 C; 642 G; 697 T; 0 U; 0 Other;

Query Match 13.9%; Score 793.8; DB 8; Length 2597;
Best Local Similarity 99.7%; Pred. No. 6,1e-181;
Matches 795; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

579 CGAGTATTTTCCACCATCTCCAGCCGGAAGTACCAAGAACTCTGAGCGGATGGCAT 638
1 CGAGTATTTTCCACCATCTCCAGCCGGAAGTACCAAGAACTCTGAGCGGATGGCAT 60
639 GTTCGGGTACGCTCTTCATGATGAGTTCGTGGCTCGATGATTAAGATCCCTTGGACAC 698
61 GTTCGGGTACGCTCTTCATGATGAGTTCGTGGCTCGATGATTAAGATCCCTTGGACAC 120
699 GTTCACCATCATCCCTGATCTTGATATCTACTATGTTTATGGTTTATAGAGTGGCAATT 758
121 CTTCACCATCATCCCTGATCTTGATATCTACTATGTTTATGGTTTATAGAGTGGCAATT 180
759 TGTCTACTTTTGAACCTTCAACCTGAGATGTTGTTCTCCACAGGCTCCACACCAAGGA 818
181 TGTCTACTTTTGAACCTTCAACCTGAGATGTTGTTCTCCACAGGCTCCACACCAAGGA 240
819 GCAGGTGTATACATCCAGCTGTGAGGCTTTGCAAGAGGACACAGCTTCAATCTCTA 878
241 GCAGGTGTATACATCCAGCTGTGAGGCTTTGCAAGAGGACACAGCTTCAATCTCTA 300
879 TGTAGAGGTGAGCCATGAGCTGTGAGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 938
301 TGTAGAGGTGAGCCATGAGCTGTGAGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 360
939 CTACCTGTCCAAAGCGGGGCGGTGTTGGCAGAGACCTTGGATTCATCCAGATGATGA 998
361 CTACCTGTCCAAAGCGGGGCGGTGTTGGCAGAGACCTTGGATTCATCCAGATGATGA 420
999 CCGTCTCTTCAACCGCTTCTCCAGAGGCGCAAGAGGCAAAATGAAATCCCTGATGATGC 1058
421 CCGTCTCTTCAACCGCTTCTCTCCAGAGGCGCAAGAGGCAAAATGAAATCCCTGATGATGC 480
1059 GGGCCCTGTGATCTTTCATCTTGAGAGCATTAATGACCGCATTAAGAGCGGCTGCAATC 1118
481 GGGCCCTGTGATCTTTCATCTTGAGAGCATTAATGACCGCATTAAGAGCGGCTGCAATC 540
1119 TTGTTACCGGGGAGGAGGAGCGCTGAGACCTGGCGCTCAAGGAGGAGGAGCATCCCTG 1178
541 TTGTTACCGGGGAGGAGGAGCGCTGAGACCTGGCGCTCAAGGAGGAGGAGCATCCCTG 600
1179 CACAGAGTGGCTCTTAACCATTAACATTAACATTAACATTAACATTAACATTAACATTAAC 1238
601 CACAGAGTGGCTCTTAACCATTAACATTAACATTAACATTAACATTAACATTAACATTAAC 660
1239 GGAAGTGTCCAGATGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1298
661 GGAAGTGTCCAGATGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 720
1299 GTCTGTATGCAATATGTCTTACAAGAACCACTCTGTGGCTTTGTGTGGACCAAAAGTGG 1358
721 GTCTGTATGCAATATGTCTTACAAGAACCACTCTGTGGCTTTGTGTGGACCAAAAGTGG 780
1359 CAAGCTGAAGAAGGTGC 1375
781 CAAGCTGAAGAAGGTGC 797

RESULT 13

ADA01231
ID ADA01231 standard; cDNA; 2597 BP.

XX ADA01231;
AC ADA01231;
XX 06-NOV-2003 (first entry)
DT Human PRO polynucleotide #46.
DE
XX

XX Human, gene; se: PRO; secreted polypeptide; transmembrane polypeptide;
XX tumour necrosis factor- α ; TNF- α ; blood; chondrocyte cell; tumour;
XX cancer; lung; colon; breast; prostate; rectum; kidney; liver;
XX microvascular endothelial cell; endothelial cell tube formation.
XX

OS Homo sapiens.

XX US2003068782-A1.

XX 10-APR-2003.

XX 16-SEP-2002; 2002US-00245851.

XX 27-APR-1999; 99US-0131271P.

XX 29-OCT-1999; 99US-0162506P.

XX 02-DEC-1999; 99WO-US028551.

XX 29-AUG-2001; 2001WO-US027099.

XX 18-JUL-2002; 2002US-00197942.

XX (GENT) GENENTECH INC.

XX Baker KP, Baton DL, Filvaroff E, Goddard A, Grimaldi JC,
XX Gurney AL, Smith V, Stephan JP, Watanabe CK, Wood WI, Zhang Z,
XX Fong S;
XX WPI: 2003-625487/59.
XX P-FSDB; ADA01232.

XX Novel isolated PRO polypeptides e.g. PRO281 and PRO1560, useful in the
XX preparation of a medicament for treating a condition responsive to PRO
XX polypeptide, and as therapeutic agents e.g. vaccines.
XX

XX Claim 2, Fig 91, 308pp; English.

XX The invention relates to isolated human PRO polypeptides (secreted and
XX transmembrane polypeptides) and the polynucleotides encoding them. The
XX invention also relates to an antibody which specifically binds to a PRO
XX polypeptide, a method for stimulating the release of tumour necrosis
XX factor- α (TNF- α) from human blood, a method for stimulating the
XX proliferation or differentiation of chondrocyte cells and a method for
XX detecting the presence of a tumour in a mammal (e.g. lung, colon, breast,
XX prostate, rectal, kidney and liver tumours). The polynucleotides are
XX useful in molecular biology, including uses as hybridisation probes, in
XX chromosome and gene mapping, in generating antisense RNA and DNA and in
XX gene therapy. The polynucleotides may also be used in preparing PRO
XX polypeptides by recombinant techniques and in generating either
XX transgenic animals or knock-out animals which are useful in the
XX development and screening of therapeutically useful reagents. The PRO
XX polypeptides or antibodies are used in preparing a medicament for
XX treating a condition responsive to the polypeptides or antibodies, such
XX as tumour, for stimulating and inhibiting proliferation of human
XX microvascular endothelial cells and for inducing endothelial cell tube
XX formation. This sequence represents a human PRO polynucleotide of the
XX invention.

XX Sequence 2597 BP; 688 A; 570 C; 642 G; 697 T; 0 U; 0 Other;

Query Match 13.9%; Score 793.8; DB 8; Length 2597;
Best Local Similarity 99.7%; Pred. No. 6,1e-181;
Matches 795; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

579 CGAGTATTTTCCACCATCTCCAGCCGGAAGTACCAAGAACTCTGAGCGGATGGCAT 638
1 CGAGTATTTTCCACCATCTCCAGCCGGAAGTACCAAGAACTCTGAGCGGATGGCAT 60

Db 121 CTTCACCATCATCCTGACTTGTATATATCTATGTTTATGTTTATGACATGGCACTT 180
 759 TGTCTACTTTTGAACCTTCCAACTGAGATGGTGTCTCACAGGCTTCACACCAAGGA 818
 Db 181 TGTCTACTTTTGAACCTTCCAACTGAGATGGTGTCTCACAGGCTTCACACCAAGGA 240
 Qy 819 GAGAGTGTATATATCCAAAGCTGTGAGGCTTTGACAGAGACACAGCTTCACTCTTA 878
 Db 241 GAGAGTGTATATATCCAAAGCTGTGAGGCTTTGACAGAGACACAGCTTCACTCTTA 300
 Qy 879 TGTAGAGTGTCCCATTTGCTGTAGAGCGAGTGGGTGAGTACCGCTGTGACAGGCTGC 938
 Db 301 TGTAGAGTGTCCCATTTGCTGTAGAGCGAGTGGGTGAGTACCGCTGTGACAGGCTGC 360
 Qy 939 CTACTGTGTCCAAAGCGGGGGCGGTGTGGAGACACCTTGGAGTCCATCCAGATGATGA 998
 Db 361 CTACTGTGTCCAAAGCGGGGGCGGTGTGGAGACACCTTGGAGTCCATCCAGATGATGA 420
 Qy 999 CTGCTCTTACCGTCTTCTCCAGAGGCGAGAAATGAATCCCTGATGATGATC 1058
 Db 421 CTGCTCTTACCGTCTTCTCCAGAGGCGAGAAATGAATCCCTGATGATGATC 480
 Qy 1059 GGGCTGTGTGATCTTCTTATGAGAGATTAATGACCGCATTTAGAGCGGCTGCAATC 1118
 Db 481 GGGCTGTGTGATCTTCTTATGAGAGATTAATGACCGCATTTAGAGCGGCTGCAATC 540
 Qy 1119 TTGTACCGGGGGGAGAGGACGCTGACCTGGCTCAAGGTGAGAGACATCCCTG 1178
 Db 541 TTGTACCGGGGGGAGAGGACGCTGACCTGGCTCAAGGTGAGAGACATCCCTG 600
 Qy 1179 CAGCAGTGGCTCTTAAACATTGACGATTAATCTTGTGGCTGACATGATGCTCCCT 1228
 Db 601 CAGCAGTGGCTCTTAAACATTGACGATTAATCTTGTGGCTGACATGATGCTCCCT 660
 Qy 1239 GGGAGTGTCCGACATGCTGGTGGAAATTCGGCTTTCACGAGAGACAGGACCGCATGAC 1298
 Db 661 GGGAGTGTCCGACATGCTGGTGGAAATTCGGCTTTCACGAGAGACAGGACCGCATGAC 720
 Qy 1299 GTCTGTATCGCATATGTCTTAAGAACCACTCTGAGCTTTGTGGGACCAAAAGTGG 1358
 Db 721 GTCTGTATCGCATATGTCTTAAGAACCACTCTGAGCTTTGTGGGACCAAAAGTGG 780
 Qy 1359 CAACTGGAAGAAGATCC 1375
 Db 781 CAACTGGAAGAAGATCC 797

RESULT 15

ADA43672
 ADA43672 standard; cDNA; 2597 BP.

ADA43672;

20-NOV-2003 (first entry)

Human cDNA encoding secreted/transmembrane polypeptide PRO34003.

ss; gene; human; PRO; secreted protein; transmembrane protein;
 endothelial cell tube formation; chondrocyte cell differentiation;
 microvascular endothelial cell; tumour; lung tumour; colon tumour;
 breast tumour; prostate tumour; rectal tumour; kidney tumour;
 liver tumour; cytoskeletal; vaccine.

Homo sapiens.

US2003073190-A1.

17-APR-2003.

09-SEP-2002; 2002US-00238283.

01-JUL-1998; 98US-0091358P.

PR 02-JUN-1999; 99NO-US012252.
 PR 20-JUL-1999; 99US-0144758P.
 PR 28-JUL-1999; 99US-0146222P.
 PR 25-AUG-1999; 99US-00380137.
 PR 30-MAR-2000; 2000WO-US008439.
 PR 02-JUN-2000; 2000WO-US015264.
 PR 29-AUG-2001; 2001WO-US027099.
 PR 18-JUL-2002; 2002US-00197942.
 PA (BETH) GENENTECH INC.
 XX Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC,
 PI Gurney AL, Smith V, Stephan JP, Watanabe CK, Wood WI, Zhang Z,
 PI Fong S;
 XX WPI: 2003-585304/55.
 DR P-PSDB; ADA43673.
 PT New isolated nucleic acid encoding a PRO polypeptide, e.g. PRO20080 or
 PT PRO21383, useful in molecular biology, chromosome and gene mapping, in
 PT generating antisense RNA and DNA, and in gene therapy.
 XX Claim 2; Fig 91; 352pp; English.
 PS The invention relates to an isolated secreted/transmembrane (PRO)
 CC polypeptide, having at least 80% sequence identity to a sequence selected
 CC from any one of the 57 amino acid sequences given in specification, or to
 CC a sequence encoded by a nucleic acid molecule selected from any one of
 CC the nucleic acids deposited under any of the ATCC accession numbers given
 CC in specification, or a sequence having at least 80% identity to PRO
 CC lacking its associated signal peptide, an extracellular domain of PRO
 CC with or without its associated signal peptide. Also included are vectors,
 CC transformed host cells, anti-PRO antibodies, the nucleic acids encoding
 CC PRO, PRO fusion proteins, inducing endothelial cell tube formation (by
 CC administering PRO281, PRO1560, PRO189, PRO4499, PRO6308, PRO6000,
 CC PRO10275, PRO21207, PRO20933 or PRO34274 polypeptide or its agonist) and
 CC an oligonucleotide probe derived from any one of the above nucleotide
 CC sequences. PRO6018 polypeptide is useful for stimulating the
 CC proliferation or differentiation of chondrocyte cells. PRO1333, PRO20080
 CC and PRO21383 polypeptides are useful for stimulating the proliferation of
 CC human microvascular endothelial cells. PRO6071, PRO487 and PRO6006
 CC polypeptides are useful for inhibiting the proliferation of human
 CC microvascular endothelial cells. PRO polypeptides are useful for
 CC detecting the presence of tumour in a mammal, including tumours of lung,
 CC colon, breast, prostate, rectal, kidney and liver. PRO281, PRO1560,
 CC PRO189, PRO4499, PRO6308, PRO6000, PRO10275, PRO21207, PRO20933 and
 CC PRO34274 polypeptides are useful for inducing endothelial cell tube
 CC formation. PRO or the antibody are useful in the preparation of a
 CC medicament for treating a condition responsive to PRO polypeptide. The
 CC oligonucleotide probes are useful for isolating genomic and cDNA
 CC nucleotide sequences, for measuring or detecting the expression of an
 CC associated gene, and as antisense probes. PRO nucleic acid is useful as a
 CC hybridisation probe, in chromosome and gene mapping, in the generation of
 CC antisense RNA and DNA, and for the preparation PRO polypeptides. PRO and
 CC PRO nucleic acid are useful as therapeutic agents, e.g. vaccines. The
 CC present sequence encodes a PRO protein.
 XX Sequence 2597 BP; 688 A; 570 C; 642 G; 697 T; 0 U; 0 Other;

Query Match 13.9%; Score 793.8; DB 8; Length 2597;

Best Local Similarity 99.7%; Pred. No. 6; 1e-181; Mismatches 2; Indels 0; Gaps 0;

Matches 795; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 579 CGAGTATTTTCCACCATCTTCCAGCCGGAACCTGACCAAGAACTCTGAGCGGATGGCAT 638
 Db 1 CGAGTATTTTCCACCATCTTCCAGCCGGAACCTGACCAAGAACTCTGAGCGGATGGCAT 60
 Qy 639 GTTCGGGTAGCTCTTCATGATAGTTCGTGGCTCGATGATTAAGATCCCTTCGACAC 698
 Db 61 GTTCGGGTAGCTCTTCATGATAGTTCGTGGCTCGATGATTAAGATCCCTTCGACAC 120
 Qy 699 CTTACCATCATCCCTGACTTGTATATCTATGTTTATGAGCATGGCACTT 758


```
Db 121 CTTCAACATCAATCCCTGACTTGTATATCTACTATGTGTATTAGCAGTGGCAACT 180
QY 759 TGTCTACTTTTGTGACCCCTGCAACTGAGTGTGTCTCCACAGGCTCCACCAAGAA 818
Db 181 TGTCTACTTTTGTGACCCCTGCAACTGAGTGTGTCTCCACAGGCTCCACCAAGAA 240
QY 819 GAGGATGATATCAATCCAGTGTGTGAGGTTTGCAGAGGAGCAAGGCTTCAACTCTTA 878
Db 241 GAGGATGATATCAATCCAGTGTGTGAGGTTTGCAGAGGAGCAAGGCTTCAACTCTTA 300
QY 879 TGTAGAGTGTCCCATTTGAGTGTGAGGCAAGTGGGTGAGTACCGCTGTGTGAGGCTGC 938
Db 301 TGTAGAGTGTCCCATTTGAGTGTGAGGCAAGTGGGTGAGTACCGCTGTGTGAGGCTGC 360
QY 939 CTAACTGTCCAAAGCGGGGCGGTGTGAGGCAAGTGGGTGAGTACCGCTGTGTGAGGCTGC 998
Db 361 CTAACTGTCCAAAGCGGGGCGGTGTGAGGCAAGTGGGTGAGTACCGCTGTGTGAGGCTGC 420
QY 999 CCTGCTCTTCAACCGTCTTCTCCAGAGGCGCAAGCGGAAATGAATCCCTGATGATC 1058
Db 421 CCTGCTCTTCAACCGTCTTCTCCAGAGGCGCAAGCGGAAATGAATCCCTGATGATC 480
QY 1059 GGGCCCTGTGATCTTATCTTGAAGGAGTAAATGACCGCATTTAAGAGCGGCTGATC 1118
Db 481 GGGCCCTGTGATCTTATCTTGAAGGAGTAAATGACCGCATTTAAGAGCGGCTGATC 540
QY 1119 TTGTACCGGGGCGAGGCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1178
Db 541 TTGTACCGGGGCGAGGCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 600
QY 1179 CAGCAGTGTGCTTTAATCAATGACATTAATCTTGTGTGCTGTGATGATGATGCTCCCT 1238
Db 601 CAGCAGTGTGCTTTAATCAATGACATTAATCTTGTGTGCTGTGATGATGATGCTCCCT 660
QY 1239 GGGAGTGTGCAATGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1298
Db 661 GGGAGTGTGCAATGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 720
QY 1299 GTCTGTCAATGCAATGTGTGCAAGAACCACTGTGTGCGCTTTGTGTGCGCAACCAAGTGTG 1358
Db 721 GTCTGTCAATGCAATGTGTGCAAGAACCACTGTGTGCGCTTTGTGTGCGCAACCAAGTGTG 780
QY 1359 CAGCTGAAGAGATCC 1375
Db 781 CAGCTGAAGAGATCC 1375
```

Search completed: February 19, 2004, 16:52:25
Job time : 1359 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 16:18:27 : Search time 259 Seconds
(without alignments)

12193.921 Million cell updates/sec

Title: US-09-964-956-12

Sequence: 1 atgaagccatgccctgaa.....gcttagacagctgaataaa 5691

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
- 2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
- 3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
- 4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
- 5: /cgn2_6/ptodata/2/ina/PCtus.COMB.seq:*
- 6: /cgn2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	DB ID	Description
1	579.2	10.2	2477	4	US-09-907-794A-169
2	579.2	10.2	2477	4	US-09-905-125A-169
3	579.2	10.2	2477	4	US-09-902-775A-169
4	284.2	5.0	3458	4	US-09-023-655-603
5	201.2	3.5	4707	3	US-09-181-706-1
6	201.2	3.5	4707	3	US-09-458-791-1
7	201.2	3.5	4707	3	US-09-458-066-1
8	201.2	3.5	4707	4	US-09-458-065-1
9	86.4	1.5	1024	4	US-09-328-475C-37
10	57.8	1.0	288	4	US-09-313-284A-6281
11	49.6	0.9	7218	1	US-08-232-463-14
12	49.4	0.9	1867	2	US-08-607-509-3
13	49.4	0.9	1867	2	US-08-634-642-3
14	49.4	0.9	1867	2	US-08-989-370-3
15	49.4	0.9	1867	4	US-09-398-169-3
16	49.2	0.9	4626	1	US-08-306-651B-22
17	49.2	0.9	4626	5	PCT-US93-06251-27
18	48.6	0.9	3012	4	US-09-551-974A-94
19	48.6	0.9	3012	4	US-09-565-501A-94
20	48.6	0.9	3012	4	US-09-639-206A-94
21	48.6	0.9	3012	4	US-09-874-923-94
22	48.6	0.9	4917	4	US-09-551-974A-100
23	48.6	0.9	4917	4	US-09-565-501A-100
24	48.6	0.9	4917	4	US-09-639-206A-100
25	48.6	0.9	4917	4	US-09-874-923-100
26	48.6	0.9	4929	4	US-09-551-974A-98
27	48.6	0.9	4929	4	US-09-565-501A-98

28	48.6	0.9	4929	4	US-09-639-206A-98	Sequence 98, Appl
29	48.6	0.9	4929	4	US-09-874-923-98	Sequence 98, Appl
30	48.2	0.8	1241	1	US-08-471-033-39	Sequence 39, Appl
31	48.2	0.8	1241	1	US-08-471-033-42	Sequence 42, Appl
32	48.2	0.8	1241	2	US-08-471-044-35	Sequence 39, Appl
33	48.2	0.8	1241	2	US-08-471-044-42	Sequence 42, Appl
34	48.2	0.8	1241	2	US-08-463-483A-39	Sequence 39, Appl
35	48.2	0.8	1241	2	US-08-463-483A-42	Sequence 42, Appl
36	48.2	0.8	1241	2	US-08-471-046A-39	Sequence 39, Appl
37	48.2	0.8	1241	2	US-08-471-046A-42	Sequence 42, Appl
38	48.2	0.8	1241	2	US-08-470-565B-39	Sequence 39, Appl
39	48.2	0.8	1241	2	US-08-470-565B-42	Sequence 42, Appl
40	48.2	0.8	1241	2	US-08-469-334-39	Sequence 39, Appl
41	48.2	0.8	1241	2	US-08-469-334-42	Sequence 42, Appl
42	48.2	0.8	1241	3	US-09-300-529-39	Sequence 39, Appl
43	48.2	0.8	1241	3	US-09-300-529-42	Sequence 42, Appl
44	48.2	0.8	1358	1	US-08-471-033-45	Sequence 45, Appl
45	48.2	0.8	1358	2	US-08-471-044-45	Sequence 45, Appl

ALIGNMENTS

RESULT 1
US-09-907-794A-169
Sequence 169, Application US/09907794A
Patent No. 6635468
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroft, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gertlisen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Guirey, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/907,794A
CURRENT FILING DATE: 2001-07-17
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547

PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 169
LENGTH: 2477
TYPE: DNA
ORGANISM: Homo sapiens
US-09-907-794A-169

Query Match 10.2%; Score 579.2; DB 4; Length 2477;
Best Local Similarity 67.1%; Pred. No. 1,4e-131;
Matches 837; Conservative 0; Mismatches 408; Indels 3; Gaps 1;

QY 151 TTCAATCACTGCTGCTGATGATGAGAGACAGACATTTACTTGGGGCCCTGATCCG 210
DB 976 TTCAACCACTTGACCGCTCCACCAAGGAGCGGGCCGCTGATGAGGGCCATCAACCG 1035
QY 211 ATTCAAGAGCTCTGACGAGCCTGAGGCTTGTGATGAGATGAGAGAGGCGGAGAG 270
DB 1036 GTCTAATAGCTGAGAGGCACTGACATCCAGGTGCTATAGAGAGGCGGAGAG 1095
QY 271 GACAAACCCCAAGTTTACCCACCCCGCATGCTCCAGACCTGCAATGAGCCCTGACAC 330
DB 1096 GACAAACAGTCTGCTTACCCCGCCCTCATGCTGAGCCCTGACAGCAAGTCTACCC 1155
QY 331 ACCAACAATGTCACAGAGATGCTCTCATAGACTTACAGAGAGAGAGGCTGATGCTGT 390
DB 1156 ACCAACAATGTCACAGAGATGCTCTCATAGACTTACAGAGAGAGAGGCTGATGCTGT 1215
QY 391 GGAAGCCTGTACCAAGGATCTGCAAGCTGCTGAGGCTGAGAGAGCTCTTCAAGCTGG 450
DB 1216 GGAAGCCTGTACCAAGGAGGCTGCTGAGGCTGCTGAGATGACCTCTTCACTGCTGG 1275
QY 451 GAGCCTATCATAGAGAGAGCACTATCTGTGAGGTGCAAGAGAGGCTGAGTCTT 510
DB 1276 GAGCCTATCCAGAGAGAGAGCACTTACTGTCCAGTGTCAAGAGAGGCGCAATGTAC 1335
QY 511 GGAAGTGTCTCTCTTACAGAGCACTGAGTGAAGCTGTTCAATGCTCCAGGAGTGT 570
DB 1336 GGAAGTGTCTCTCTTACAGAGCACTGAGTGAAGCTGTTCAATGCTCCAGGAGTGT 1395
QY 571 GGAAGACCCGAGTATTTTCCAGCATCTCCAGCCGAGAACTGACCAAGAACTGAGGG 630
DB 1396 GGAAGACCCGAGTATTTTCCAGCATCTCCAGCCGAGAACTGACCCGAGAACTGAGTC 1455
QY 631 GATGAGATGCTGCTGCTTACAGCACTGAGTGAAGTGTGAGCTGATGATTAAGTCC 690
DB 1456 TCAGCAATGCTGCACTATGAGCTACAGAGGATTTTGTCTCTCTCATCAAGATCC 1515
QY 691 TCAGCAATGCTGCACTATGAGCTTGAATGATGATGCTGATGATGATGATGATGAT 750
DB 1516 TCAGCAATGCTGAGCTGCTGCTTCCAGCTTGAATGATGATGATGATGATGATGAT 1575
QY 751 GGAACAATGCTGCTGCTTGAACCTGCAAGCTGAGA---TGATGCTGCAAGAGGCTCC 807
DB 1576 GGAAGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1635

QY 808 ACCAACAAGAGAGGCTGATATACATCCAGCTGCTGAGGCTTTCAGAGAGACAGACC 867
DB 1636 TCAGCTGAGAGCTCTTCTACATCCAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1695
QY 868 TTCAATCTCTATGATGAGAGTCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 927
DB 1696 TTCAATCTCTATGATGAGAGTCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1755
QY 928 CTGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 987
DB 1756 CTGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1815
QY 988 CCAGATGATGACCTGCTCTTCAACCGCTTCTCCAGAGGCGGAGAGGAGAAATGAAATCC 1047
DB 1816 AGCAGAGAGATGATGATGCTTCTTCCATCTTCTCCAGAGGCGGAGAGGAGATGAT 1875
QY 1048 CTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1107
DB 1876 CCAGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1935
QY 1108 CGGCTGAGTCTGCTTTCACCGGCGGAGGAGAGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1167
DB 1936 CGGCTGAGTCTGCTTTCACCGGCGGAGGAGAGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1995
QY 1168 GACATCCCTGCTGAGAGTGGCTCTTAACTTGAAGATGATGATGATGATGATGATGATG 1227
DB 1996 GACGCTGAGTCAAGAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2055
QY 1228 AATGCTCCCTGAGAGTGTCCGAGATGATGATGATGATGATGATGATGATGATGATG 1287
DB 2056 AACAGAGCTGAGAGGCTGATCACTGAGAGGCTGAGCTGATGATGATGATGATGATGAT 2115
QY 1288 GACGATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1347
DB 2116 GACGATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2175
QY 1348 ACCAAGTGGCAAGCTGAGAGAGATCCGGGTGATGAGACCCAGAGGCT 1395
DB 2176 ACTAAGATGCGAAGCTGAGAGAGATCCGGGTGATGAGATGATGATGATGATGATG 2223

RESULT 2
US-09-905-125A-169
Sequence 169, Application US/0905125A
Patent No. 6664376
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Flivarov, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerlitsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Grimaldi, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kijavini, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerltzen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Grimaldi, Paul J.
APPLICANT: Guirney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Proul, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mckey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/902,775A
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 169
LENGTH: 2477
TYPE: DNA
ORGANISM: Homo sapiens
US-09-902-775A-169

Query Match 10.2%; Score 579.2; DB 4; Length 2477;
Best Local Similarity 67.1%; Freq. No. 1.4e-13;
Matches 837; Conservative 0; Mismatches 408; Indels 3; Gaps 1;
151 TTCAATCACTGTGTGATGAGAGACAGACATTACTTGTGGGCGCGTCAATCGG 210
976 TTCAACCACTGTGACCGTCCACCAAGAGAGCGGGCGCGTCTATGTGGGCGCATCAACCGG 1035
211 ATTCAAGCTCTCCAGGACCTGAAGCTTGTGTGACGATGAGACAGGCGCGAGAG 270

1036 GTCTATTAAGCTGTACAGGCAACCTGACATTCAGATGGCTCTATTAAGACAGGGCGAAGAG 1095
271 GACAACCCCAAGTGTATACCAACCCCGCATGTGTCAGATGAGACCCCTGACAC 330
1096 GACACAAAGTCTGTATACCGCCCTCATGTGACAGCCCTGACAGCAAGTGTCAACCTTC 1155
331 ACCAACAATGTCAAAAGATGCTCCATAGATCAAGAGAAAGAGCTGATGGCTGT 390
1156 ACCAACAATGTCAAAAGCTGTCTCATATGATCTGTGAGAACCGCTGTGGCTGT 1215
391 GGGAGCTGTACCAAGGATGTGCAAGCTGTGAGGCTGTGAGAACCTTTCAAGCTGGG 450
1216 GGGAGCTGTACCAAGGAGGCTGTGAGGCTGTGAGAACCTTTCAAGCTGGG 1275
451 GAGCTTATATAGAAAGAGGACCTATGCTGAGTGTCAAGAGAGGCTGAGTCTT 510
1276 GAGCTTATATAGAAAGAGGACCTATGCTGAGTGTCAAGAGAGGCTGAGTCTT 1335
511 GAGTGTATGTCTCTTACCAAGCACTGATGACAGCTGTTCATGTCACGAGTGTAT 570
1336 GGGGTATGTGGCTGTGAGGATGAGATGAGCAAGCTTTCATGCGCTGTGAT 1395
571 GGGAGCCCGAGTATTTTCCACCATTCAGCGCGAACTGACCAAGACTGTGAGGCG 630
1396 GGGAGGAGGATTAATTCCTGACCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1455
631 GATGAGTGTGTGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 690
1456 TCAGCATTGTGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1515
691 TCGACACCTTACCAATATCTGATGATGATGATGATGATGATGATGATGATGATGAT 750
1516 TCAGCATTGTGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1575
751 GGCACCTTGTGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 807
1576 GGGGCTTGTGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1635
808 ACCACCAAGAGGAGTGTGATATCAAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 867
1636 TCCCTGTGAGGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1695
868 TTCAACTCTGTATGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 927
1696 TTCAACTCTGTATGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1755
928 CTGACAGTGTGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 987
1756 CTGACAGTGTGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1815
988 CCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1047
1816 ACCACCAAGAGGAGTGTGATATCAAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1875
1048 CTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1107
1876 CCGCATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1935
1108 CGGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1167
1936 CGGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1995
1168 GACATCCCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1227
1996 GACATCCCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2055
1228 ATGTCTCCCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1287
2056 AACCAAGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2115
1288 GACCCATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1347

Mon Feb 23 09:34:02 2004

us-09-964-956-12.rn1

Page 6

NAME: Henry, Janis C
REGISTRATION NUMBER: 34,347
REFERENCE/DOCKET NUMBER: 2631-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)470-4189
TELEFAX: (206)233-0644
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4707 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOHETICAL: NO
FEATURE:
ANTI-SENSE: NO
NAME/KEY: CDS
LOCATION: 1..4707
US-09-181-706-1

Query Match 3.5%; Score 201.2; DB 3; Length 4707;
Best Local Similarity 49.1%; Pred. No. 4.3e-39;
Matches 824; Conservative 0; Mismatches 728; Indels 126; Gaps 6;

3935 TTCCGCTCTGAGACTATGAACTTACACCATGCGGCTGCTGTCCTCCAGAAATTGAAGACC 3994
3038 TTCCCTCTTCTTACTACCAACATTTTGTCTGAGAACTTCTTCCCTGAGTCAGGTGGCT 3097
3995 ACCCTGCTCCGGGACCTTGAAGTCCCGGCTACCGGAGAGCGGTGTGAAGAGCC 4054
3098 TCACCCACATCTTCACTGAGATATGCAATACAGAGAGCCCAACAGATGAATATC 3157
4055 TGAAGCTCTGCGCCAGCTACCAACCAAGGTCTGCTGCTTCTTCAATCCGACAGC 4114
3158 TCACAGCTTTGATGCTTAACTTCTGATTAAGATCTTCTTCTTCTTCTTCTTCTTCA 3217
4115 TTGAGTCCAGCGTATGCTTCTCATGCGGACCGTGGCAACGTGCTCACTCATATGA 4174
3218 TTGAAAGAGAGAACTTTCTGTGAAGAGACAGGTGTCTTCTTCTTCTTCTTCTTCA 3277
4175 CCGTGTGAGAGAGAGCTGAGATGCGCACTGATGTGTGAAGAGAGCGTGGCCAGC 4234
3278 TTGCACTGAAACCAAGCTGTCTTCACTGACAGCATCTAAGGTGCTGACAGGAGCT 3337
4235 TCATTGACAAAGACTGAGAGACCAACCACTTAAAGTGTCTGCTGAGAGAGCTGAGT 4294
3338 TGAAT-----GGAAAGTGTATGATACATGACCCGAACTCATGCTGAGAGCGAGT 3391
4295 CAGTGTGAGAAAGATGCTGACCAATGTTTACTTCTCTCTCAAAAGTTCCCTCAGG 4354
3392 CCGTCTGCAAAAACCTCTCAACAACTGATGTCTGCTCTTCTTCTTCTTCTTCTTCTG 3451
4355 AGTGTCTGGGAGCCCTCTTCTCTCTGTTCTGTGTCATCAAGAGAGAGATGAAGAG 4414
3452 AGACTGTGGAGAGCCCTTCTTATTTGCTGTGAGCACTGAAACAGAAATTAACAAGG 3511
4415 GCCCACTGAGCGCATGAGGGGAGGCCGCTACTCTCTTGAAGAGAGAGAGCTCAACC 4474
3512 GTCCCGTGTATGATCACTTGTGAAGCCCTGTACACACTTATGAGAGCTGGCTGTGT 3571
4475 GCCAGAGATTGACTCAAAACCTGTCTGAGCTGTGTCAAGCCAGAGATGCAACA 4534
3572 GCGAGGTCGGGAATTCAGTCTGTGCAATTAACCTCTCTTGAATAAATCCCGGAAA 3631
4535 GC-----CCGAGGTCCAGTAAAGTCTCAACTGTGACACATCA 4576
3632 ACAGAGTGCAGATGTCTGTGGAATATTTCACTCATATTTCTGACGTGTGACACATTTG 3691
4577 CTGAGGTCAAGAGAGATCTGATGCACTTCAAGATGAGCTTCTCCACCGGC 4636
3692 GCCAAGCCAAAGAAATTTTCCAGCATTTTAAGCAAAATGCTCTCTTATGAGC 3751
4637 CCAAGCTGCAGATGATCTGAGTGGCAGACAGAAAGTGGGGCAGAGATGATCTTGC 4696

3752 TTCCAGCTTATGAATATGCTCTTGAAGCTTCAATATGGGACACAGCAAGAAAGACTTCTGG 3811
4697 AGAGTGAAGACATCAACCAAGATTTGAATGATTTGAAGGAGCTGAACACTGCGCC 4756
3812 ACATGACAGTTCCTCCGATTTCTTGAAGATGAATCAACAGCTAAACACCATTTGGCC 3871
4757 ACTACAGGTGCAATGATTTCCGTGTGCACTTGAAGTCCAGAGAGTGAAGCTTATA 4816
3872 ACTATGAGTATCAATGATGATCACTTAAAGTCTTTAAGAGATGACAAATTTTACTT 3931
4817 ACCAGTGAACAACCTCAACCGTCTCAGAGCTCAGCAAGTAAATATGAAACATGATCC 4876
3932 CAGATGTGAGTATCT-----GAGTACCATCTGCCATTTGATTTTA----- 3972
4877 GGTACAGGGGAGCCCGACAGCTTCGCTACAGACACTTATGATCTCTGACCTGG 4936
3973 ----- 3972
4937 AGAGTGAAGTCAAGATGTGCACTAGTGAAGAACACAGAGACGAGACAGAGAG 4996
3973 -----CGAGATTCGAGAGCATTCAGATGTGAGAGAAAGAGC-----ATC 4015
4997 GGGACCGGGGAGCAAGATGTGTCTGAATATCTACTGACCCGACTCTGCGCACTAAG 5056
4016 GAGGGAAGCAAGTCAAGTAAAGAAATATGATCTGCAAGCTGCTGTGACCAAG 4075
5057 GCACATGCAAGATTTGTGATGACCTCTTGAAGCATCTTCAAGCAGAGACAGCCGTG 5116
4076 TGGCAATCTATCTGTGCTGAAAACTTTTGAAGCATTTGAGTTTACCAACAGCA 4135
5117 GCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5176
4136 GACT-----CATTTGCTATTAATATCTTTTGTGACTTTTGTGAGCGCCAGGCTGAA 4189
5177 AAGATGCAATTCATGACCGGACGCTGCGCATACCTGGAAGAGCAATTCCTGCGCCCTGA 5236
4190 ACAAATAATCAAGATCTGAGCTGTAATTTTGAAGAAACAAACAGCTTCTCTTC 4249
5237 GGTCTTGGTCAATGATCAAGAACCGGAGTTGTGTTGATCATCAATGAAGAGCA 5296
4250 GCTTGGGTAATCATCTGGAAGACCTCACTTGTCTTGTGACTTAAGAGACACAC 4309
5297 TCACAGCGCTGCTCTCTGTGTGTCTGAGACCTTCAAGCATCTTGTCTCAAGTGA 5356
4310 ATATAGCGGCTTTGTGATGATGCTGAGCAATTCATGATGATTTCTCTCAGC 4369
5357 AGCAGCGCTGGGCAAGACTCGCTCTCAACAGCTGTGATGCGAAGAGATCCCA 5416
4370 AGCAGCAATGAGGAGAGAGACCAACTAATTAAGCTCTCTATGCGAAGATATCCAA 4429
5417 GCTTCAAGATTTGGGTGAGAGATTTATCTCAACATAGGAGAGATGCCAGCATAGG 5476
4430 CTTCAAAAGAGAGTAAATCTTATTTACAAACATACAGGATTTGCTCTCAT 4489
5477 ACCAAGATGAAGCATACCTGCTGAGAGTCCCGATGCAATGAATGAGTTCAACA 5536
4490 CTTCAAGAAATGAGAAATTTTAACTCAGGAATCTAAGAAACATGAATAATTAATG 4549
5537 CCATGAGTCACTCTCAGAGATCTTCTCTATGTGCGAAATACAGCGAGATCT 5594
4550 AAGAGTGGCTTGAAGAAATTTAACAATATGTAATTAATTTTGTGATGATCT 4607

RESULT 6
US-09-458-791-1
Sequence 1, Application US/09458791
Patent No. 6174689
GENERAL INFORMATION:
APPLICANT: Spilgers, Melanie
TITLE OF INVENTION: VIRAL ENCODED SENAPHORIN PROTEIN
RECEPTOR DNA AND POLYPEPTIDES
NUMBER OF SEQUENCES: 10

Mon Feb 23 09:34:02 2004

us-09-964-956-12.rml

Page 7

CORRESPONDENCE ADDRESS:
ADDRESSEE: Janis C. Henry
STREET: 51 University St.
CITY: Seattle
STATE: WA
COUNTRY: US
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS/Windows 95
SOFTWARE: Word for Windows 95, 7.0a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/458.791
FILING DATE: 10-Dec-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/958.598
FILING DATE: 28-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Henry, Janis C
REGISTRATION NUMBER: 34,347
REFERENCE/DOCKET NUMBER: 2631
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)470-4189
TELEFAX: (206)233-0644
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4707 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..4707
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-458-791-1
Query Match 3.5%; Score 201.2; DB: 3; Length 4707;
Best Local Similarity 49.1%; Pred. No. 4.3e-39;
Matches 824; Conservative 0; Mismatches 728; Indels 126; Gaps 6;

QY 4355 AGTGTGTGGGAGGCCCTCTTCTCCCTGTTCTGTGTCATCAAGACAGATGAGAGG 4414
DB 3452 AGACTGTGGAGAGGCCCTTCTATTGCTGTGAGAGCTGTGACCAAGAAATTAAACAAG 3511
QY 4415 GCCCATTTGAGCCATCAAGGAGGCGGCTACTCTTTAGAGGAGCAAGCTCATCC 4474
DB 3512 GTCCCGTAGATGTATCACTTGCAAAAGCCCTGTACACATTAAAGAACTGCGTGTGT 3571
QY 4475 GCCAGCATTTGACTACAAACCCCTGTGCTGAGCTGTGACGCCAGACATGCCAACA 4534
DB 3572 GCGAGGTTCCGGAATTCACTGATCTGTGGATTAAGTCGTCTTTGAAAAAATCCCGGAA 3631
QY 4535 GC-----CCCGAGTCCAGTAAAGATCCTCACTGTGACACATCA 4576
DB 3632 ACGAGGTGAGAGTGTCTGTGGAATATTCAATCAATGTTCTGACGTGACACCATTG 3691
QY 4577 CTCAAGTCAAGAGAGATTCGTGATGCCATCTTCAAGATGTGCTGCCACCGGC 4636
DB 3692 GCCAAGCCAAAGAAAGATTTTCCAGACATCTTAAGCAAAATGGCTCTCTTATGAGC 3751
QY 4637 CCAAAGCTGCAAGATGTGATCTGAGTGGCGACAAAGAGTGGGCAAGATGATCTTGC 4696
DB 3752 TTCACTTAATGAATTTGTCTTGAGCTTCAATGGGCAACGACAGAAAGACTTCTCG 3811
QY 4697 AGGATGAAGACATCAACCAAGATTGAGATTTGGAAGCGACTGAACACACTGGCCC 4756
DB 3812 ACATGCAAGTTCCTCCGTGATCTTGAAGATGAATCAACAGTTAAACCAATTGGCC 3871
QY 4757 ACTACAGATGCGAGATGTTCCGTGTGTGAGTATGATCCAGAGGTGAACAGCTATA 4816
DB 3872 ACTATGAGATTAATAATGATTCATTAAGATCTTTAAGAGATAGCAATTTACTT 3931
QY 4817 ACGCAGTGAACAATCCACCGCTTCCAGGACTTCAAGCAATTAATAAATCATGATCC 4876
DB 3932 CAGATGTGAGATATCC-----GAGTACACACTGCAATTTGATTTA----- 3972
QY 4877 GGTACACGGGAGCCCGACAGCTCCGCTCAGGACACTATGATCATCTCTGACCTGG 4936
DB 3973 ----- 3972
QY 4937 AGAGTGAAGTCAAGTGTGCACTAGTGAAGAACACAGACAGAGACCAAGAGG 4996
DB 3973 -----CGAGTTTGGAAGATTCGCAAGATGTGCAAGAAAGAAC-----ATC 4015
QY 4997 GGGACCGGGGGAGCAAGATGTGTCTGAATCTACTGACCCGACTCTGGCCACTAAG 5056
DB 4016 GAGGAGACCAAGATTCAAAGTAAAGAAATGTATGTACAAAGCTGTGTGACCAAG 4075
QY 5057 GCAACTGCAGAAATTGTGTGATGACTCTTTGAGACCATTTTCAACGCGACACCGTG 5116
DB 4076 TGGCAATTCATCTGTGCTTGAAGAACTTTTGAAGCATTTGAGTTTACCAACAGCA 4135
QY 5117 GCTTGCCCTGCCCTGCGCATCAAGTACATGTTTGACTTCTCGATGAGCAGGCTGATA 5176
DB 4136 GAGCT-----CATTTGCTAATAAATCTTTTGAATCTTTTGAAGCCGACGCTGAAA 4189
QY 5177 AACATGGCATTCATGACCGGACGTCGCACTACCTGTGAAGACAAATTGCTGCCCTGA 5236
DB 4190 ACAAATAATCAAGATCTGACGTCGACATATTGGAAGAAACAAAGCCTTCTCTTC 4249
QY 5237 GGTTTTGGGTCAATGATCAAGAACCGGAGTTTGTCTTGAACATCATTAAGAACACA 5296
DB 4250 GCTTCTGGTAAACATCTGGAAGACCTCAGTTTGTCTTGAACATTAAGAACACAAC 4309
QY 5297 TCACAGAGCGCTGCTCTGTGTGTGCTCAGACCTTCAAGACTTGTCTCAAGCTGAG 5356
DB 4310 AATATAGAGGCTGTTTGTGAGATTTGCCAGGCACTCATGATGCAATTTCTCTCAAG 4369
QY 5357 AGCACCGGCTGGGCAAGACATCGCCCTCAACAGCTCTGTATGCGCAAGGACATCCCA 5416
DB 4370 AGCAGCACTAGGAGAGAGACCACTAATTAATAGCTCTCTATGCAAGGATATCCCA 4429
QY 5417 GCTACAGAAATTGGGTGAGAGGTATTACTCAGACTAAGGAGATGCGACGATCAAGC 5476

Db 4430 CCTACAAAGAGAGATTAATCTTATTTACAAAGCATCGAGATTTGCTCCATTTGCAT 4489
Qy 5477 ACCAAGACATGAAACGATACCTGCTGAGCAGTCCCGATGCATGAAATGATTCACA 5536
Db 4490 CCTCAAAATGGAAGATTTTACTCAGAAATCTAAGAAACATGAAATGATTTAATG 4549
Qy 5537 CCAATGAGTCACTCAGAGATCTTCTCTATGTGGCAATACAGCGAGAGATCCT 5594
Db 4550 AAGAAGTGGCTTGAAGAAATTTACAAATCATCTGAAATTTTGTATGATTTCT 4607

RESULT 7
US-09-459-066-1
Sequence 1, Application US/09459066
Patent No. 6187909
GENERAL INFORMATION:
APPLICANT: Spritz, Melanie
TITLE OF INVENTION: VIRAL ENCODED SEMAPHORIN PROTEIN
TITLE OF INVENTION: RECEPTOR DNA AND POLYPEPTIDES
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Janis C. Henry
STREET: 51 University St.
CITY: Seattle
STATE: WA
COUNTRY: US
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS/Windows 95
SOFTWARE: word for windows 95, 7.0a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/459,066
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/958,598
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Henry, Janis C
REGISTRATION NUMBER: 34,347
REFERENCE/DOCKET NUMBER: 2631
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)470-4189
TELEFAX: (206)233-0644
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4707 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..4707
US-09-459-066-1

Query Match 3.5%; Score 201.2; DB 3; Length 4707;
Best Local Similarity 49.1%; Pred. No. 4.3e-39;
Matches 824; Conservative 0; Mismatches 728; Indels 126; Gaps 6;

Qy 3335 TTCCGTCCTGGAATATGAACTTACACGATGGCGGTGTTCCACGAAATGAGAC 3994
Db 3038 TTCCCTTCCTGACTAACAACATTTTCTCTGAGAACTTTCTTCTCTGAGTCAGGTGGCT 3097
Qy 3395 ACCCTGTCTCCGGAGCTTGAAGTCCGGGCTACCGGACGAGCGGTGAGAAAGCC 4054
Db 3098 TCACCAACATCTTCACTGAAGATATGATTAACAGAGACGCCAACAGAAATGAAGATC 3157

Qy 4055 TGAAGCTCTGGCCAGCTCATCAACAAGGTGTCTGCTGCTCTCATCCGACGC 4114
Db 3158 TCACAGTTTGGATGAGCCCTATCTGTATATTAAGCTTTCTTGTATCTGATCATCAACCC 3217
Qy 4115 TTGATGCCAGCGATGTTCTTCATGCGGACCGTGGCAACGGTGGCTTCACTCATATA 4174
Db 3218 TTGAAAGCAAGAACTTTTCTGTGAAGACAGGTGTCTGTTGCTCTCTTCAACA 3277
Qy 4175 CCGTGTGACAGCAAGCTGAGTACGCCCTGATGTGTGAAGCAGCTGTGGCCGACC 4234
Db 3278 TTGACTGTCAACCAAGCTGTCTTACTGTACAGCATCTTGAAGGTGTGACCAAGGACT 3337
Qy 4235 TCATTGCAAGAACTTGAGAGCAAGAACCACTTAAGTCTGCTCAGAGAGACTGAT 4294
Db 3338 TGAT-----GGAACAGTGTATGATGATGAGCGCAAACTGATGAGACGCAAGAT 3391
Qy 4295 CAGTGTGAGAGAGATGCTGACCAATGTTTACTTCTCTCTCAAAATTTCTCAAG 4354
Db 3392 CCGTGTGAGAAACTCTTCAAACTGATGTCCGTCTGCTTTCTGATTTCTCCGG 3451
Qy 4355 AGTGTCTGGGAGCCCT 4414
Db 3452 AGACTGTGAGAGCCCTTCTATTTGTGTGTGAGACTGTGAACCGAAATTAAGAG 3511
Qy 4415 GCCCATTTAGCCATGACGGGCGAGCCCGCTACTCTTGAAGCGAGGACCAAGCTATCC 4474
Db 3512 GTCCGTGATGATATCATCTTGCAAGCCCTGTACACATTAATGAAGCTGCTGTGT 3571
Qy 4475 GCCCAGATTGACTACAAACCCTGTCTCTGAGCTGTGTCAAGCCAGACATGCAAC 4534
Db 3572 GCGAGTTCGGGAATTCATGATCTGTGCATTAACGTCCTTGAATAATCCCGAAA 3631
Qy 4535 GC-----CCGAGTCCCATTAAGATCCCACTGTGACACATCA 4576
Db 3632 ACGAGATGAGATGTGTGTGGAATTTCAAGTCAATGTCTGATGTGACACCAATG 3691
Qy 4577 CTCAGTCAAGAGAAATCTGATGCCATCTTCAAGAAATGTGCTTCCACCGGC 4636
Db 3692 GCCAAGCCAAAGAAATTTTCCAGCATTTCTTAAGCAAAATGGCTCTTATAGAC 3751
Qy 4637 CCAAGCTGCATATGATGTGAGTGGCGCAAGAAATGGGGCAAGATGATCTGC 4696
Db 3752 TTCACTTAATGAATGTGTGAGCTTGAATGGCAACAGCAAGAAATCTTGTG 3811
Qy 4697 AGATGAAGATCATCACCAACCAATGAGATGTGAAGCACTGACACTGACCC 4756
Db 3812 ACATGACAGTTCCTCCGTGATTTCTTAAGATGGAATCAACAGCTTAACCATTTGGCC 3871
Qy 4757 ACTACAGGTGCGAGATGTTCGTGTGATGATGATGATGATGATGATGATGATGATGAT 4816
Db 3872 ACTATGAGATATCAAAATGATTCCTAATAAGCTTTTAAGAAATGCAAAATTTTACTT 3931
Qy 4817 ACGAGTGAACAACCTCCAGCGTCTCCAGACCTCAGCAAGTAAATATGAAATGATGCC 4876
Db 3932 CAGATGTGATCTC-----GATGACCACTGCAATTTATTTTA----- 3972
Qy 4877 GGTACAGCGAGCCCGGACAGCTCCGCTTACGGAACCTATGATCATCTTGAAGCTGG 4936
Db 3973 ----- 3972
Qy 4937 AGATGAGTGAAGATGTGAGACTAGTGAAGAACAGACGAGACGAGACGAGAGAG 4996
Db 3973 -----COAGTTGGAAGATTCAGAGATGTGAGAGAGAGAC-----ATC 4015
Qy 4997 GCGACCGGAGGAGCAAGTGTGTCTGAATCTACCTGACCCGACTCTGTGCACTAAG 5056
Db 4016 GAGGAGACCAAGTTCAAATGAAGAAATGTATCTGACAAACCTGTGTGACCAAG 4075
Qy 5057 GCACTGTGCAAGTTGTGTGATGACCTCTTGAAGCAATTTGAGACGAGCAACCGTG 5116
Db 4076 TGGCAATTCATTTGTGTGATTAAGAACTTTTAAGACATTTGAGATTACCAAGCA 4155
Qy 5117 GCTGTCCCTGCCCCCTGCGCATCAAGTATCATGTTTGAATCTTGTGATGAGCAGCTGATA 5176

```

Db      4136 GAGCT-----CCATTGCTATTAATACTTTTGAAGCCCGAGGCTGAAA 4189
Qy      5177 AACATGCGATTTCATGACCCGCGCATCTGCAAGAGCAATTCGCTGCCCTGA 5236
Db      4190 ACAAATAATCAAGATCCGAGCTGTACATTTTGGAAAAAACAACGCTTCCTTC 4249
Qy      5237 GGTTCGCTCAACATGATCAAGAACCCGAGTTTGTGTTGATCCATTAAGACAGA 5296
Db      4250 GCTTCCTGGTAAACATCTGTAAGAACCCCTCAGTTTGTGATTAAGAACACAC 4309
Qy      5297 TCACAGACGCTGCTCTCTGTGTGCTCAGACCTTCATGACCTCTGCTCAGCTCAG 5356
Db      4310 ATATGACGCTGTTTGTCTGATGTTGCCAGGCACTTCATGATGATTTTCTCAG 4369
Qy      5357 AGCACCGCTGGGCGAGAGCTCGCCCTCCAAAGCTGCTATGCCAAGACATCCCA 5416
Db      4370 AGCAGCAACTAGGAGGAGAGAGCAACAATAATAGCTTCTATGCCAAGATATCCAA 4429
Qy      5417 GCTACAGATTTGGGTGGAGAGGATTAATCTCAGACATAGGAGAGATGCCAGCATAGCG 5476
Db      4430 CTTACAAAGAGAAATTAATTTTATTACAAAGCATCAGGATTTGCTCCATTTGCTAT 4489
Qy      5477 ACCAAGCATGACGATACCTGCTGAGCAGTCCCGGATGCAATGATGATTGACA 5536
Db      4490 CCTCAGAAATGGAAGATTTTAACTCAGGATCTAAGAAACATGAATAATTAATG 4549
Qy      5537 CCATGATGCACTCTCAGAGATCTTCTCTATGTGGCAATACAGGAGAGATGCT 5594
Db      4550 AAGAGTGCGCTTGACAGAAATTTACAAATPACATGTAATAATTTTGTGATGATTC 4607

```

RESULT 8

US-09-459-065-1

Sequence 1, Application US/09459065

Patent No. 6562949

GENERAL INFORMATION:

APPLICANT: Spilgers, Melanie

TITLE OF INVENTION: VIRAL ENCODED SEMAPHORIN PROTEIN

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS: 10

ADDRESS: Janis C. Henry

STREET: 51 University St.

CITY: Seattle

STATE: WA

COUNTRY: US

ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: MS-DOS/Windows 95

SOFTWARE: Word for Windows 95, 7.0a

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/459,065

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/958,598

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Henry, Janis C

REGISTRATION NUMBER: 34,347

REFERENCE/DOCKET NUMBER: 2631

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206)470-4189

TELEFAX: (206)233-0644

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4707 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

```

MOLECULE TYPE: CDNA
HYPOTHEICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..4707
US-09-459-065-1

```

Query Match 3.5%; Score 201.2; DB 4; Length 4707;
 Best Local Similarity 49.1%; Pred. No. 4.3e-39;
 Matches 824; Conservative 0; Mismatches 728; Indels 126; Gaps 6;

```

Qy      3935 TTCGCTTCTGAGTATTAAGATTAACCATGCGGCTGCTGTTCCAGGAATTAAGACC 3994
Db      3038 TTCCCTTCCTTGACTACAAACATTTTGTGCTGAGAACTTCTTCCAGATCAGGTGCT 3097
Qy      3995 ACCCTGCTCCGGAACCTTGAAGTCCCGGCTACCGGAGAGAGCTGTGAGAAAGGCC 4054
Db      3098 TCACCCATCTTCACTGAAGATATGATTAACAGAGACGCCAAGACAGATGAAGTC 3157
Qy      4055 TGAAGCTCTTGCCAGCTCATCAACAAGAGTGTCTGCTGCTCTTCAATCCGACGC 4114
Db      3158 TCACAGCTTTGAGATGCCCTATCTGTATTAAGACTTTCTTTACTGTATCCACACCC 3217
Qy      4115 TTGAGTCCAGCGTAGCTTCTTCATGCGCGACCGTGGACAGTGGCTCACTCATGA 4174
Db      3218 TTGAAGACAGAAAGACTTTTCTGTGAAGGACAGGTGTCTGCTTCCCTTAACCA 3277
Qy      4175 CCGTGTCCAGAGAGAGCTGAGTACGCCCACTGATGTGCTGAGCAGCTGCGCCGACC 4234
Db      3278 TTGACCTGAAACAGAGTGTCTTACCTGACCAAGATCTTGAAGTGTGACAGGACT 3337
Qy      4235 TCATTGACAGAACCTGAGAGCAAGAACCAACCTTAAGCTGCTGCTCAGAGACTGAGT 4294
Db      3338 TGAT-----GAAACAGTGTAGTAACATGACGCCAAATCATGCTGAGACAGGAGT 3391
Qy      4295 CAGTGGCTGAGAGAGATGCTGACCAATGCTTACCTTCTGCTGACAGTTCCTCAAG 4354
Db      3392 CCGTGTGAAAGAACTCTCAAACTGATGATGATGATGATGATGATGATGATGATGAT 3451
Qy      4355 AGTGTCTGAGGAGGAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4414
Db      3452 AGCTGTGAGAGAGGCTTCTTATTTGCTGTGACGACTGAAACAGAAATTAACAAG 3511
Qy      4415 GCCCATTTGAGCGCATACAGGCGGAGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4474
Db      3512 GTCCGCTGAGATGATCACTTGAAGAGGCTGTACACTTAATGAACATGCGCTGTTGT 3571
Qy      4475 GCCAGCATTTGACTACAAAACCTGTCTCTGAGCTGTGAGCCGACCAATGCAACA 4534
Db      3572 GGCAGGTTCCGGAATTCAGTACTGTGCAATTAACGTGCTTTGAAATAATCCGGA 3631
Qy      4535 GC-----CCGAGGCTCCAGTAAATCTCAATCTCAATCTCAATCTCAATCTCAAT 4594
Db      3632 ACAGAGTGACAGATGCTGTGGAATATTTCACTGAATGTTCTGACTGTGACACATTTG 3691
Qy      4577 CTCAGGTCAAGAGAGATTCGTGATGCCATCTTCAAGAAATGTCCTTGTCTCCACCG 4636
Db      3692 GCCAGCCAAAGAAATTTTCCAGATTTCTTAAGAAAATGCGCTCTCTTAATGAC 3751
Qy      4637 CCAAAGCTGACATATGATTCGATGAGCGGCAAGAGAGAGAGATGATCTTGC 4696
Db      3752 TTCAGCTTAATGAATTTGCTGTGAGCTTCAATGAGCAACGACAGAAAGACTTCTG 3811
Qy      4697 AGAGTGAAGACATCAACCAAGATTGAGATTTGAAGGACCTGACACATGAGCC 4756
Db      3812 ACATGACAGTCTCTCCGATCTTGAATGAGATGAGATGAGATGAGATGAGATGAG 3871
Qy      4757 ACTACAGGTGCAAGATGCTGCTGAGCTTGAATGATGATGATGATGATGATGATGAT 4816
Db      3872 ACTATGAGATATCAATGATGATCACTATTAAGATTTTAAGAGATTAACAAATTTTACT 3931
Qy      4817 AGCAGTGAACAATCCACCGCTCTCAGGACCTCAGCAAGTAAATATGAATAACATGATCC 4876

```

```

Db      3932  CAGATGTGAGTACTTC-----GGATGACCATGTCATTGATTTTA----- 3972
Qy      4877  GGTACAGGGGAGGCCCGACAGCTCCGGCTCAGGAGACCTATGATCACTCTGACCTGG 4936
Db      3973  ----- 3972
Qy      4937  AGAGTGAAGTCAAGATGTGSCACCTAGTGAAGAACACGAGCACGAGACCAAGAGAGG 4996
Db      3973  -----CCAGATTGGAGAAATTCACAGATGTGCGAAGAAAGAGAC-----ATC 4015
Qy      4997  GGGAGCCGGGGAGGAGAGTGTGTCTGAATCTGACCTGACCCGACTCCTGGCCATAAG 5056
Db      4016  GAGGAGACACAACTTCAAAGTAAAGAAATGTATCTGACAAAGCTGTCTGACCAAG 4075
Qy      5057  GCAACCTGCAGAAAGTTGTGATGAACCTCTTGAACCACTTTCAGACCGCACACCTGG 5116
Db      4076  TGGCAATTCATTCTGTGCTTGAATAAATTGTAAGACATTTGAGATTGAGATTACCAACAGCA 4135
Qy      5117  GCTGTGCCCCCTGCCCCGACATCAAGTACATGTTGATCTTCTGATGAGAGAGGCTGATA 5176
Db      4136  GAGCT-----CCATTGCTATTAATACCTTTTGTGACTTTTGAAGCCCGAGGCTGAAA 4189
Qy      5177  AACATGGAATTCATGACCCGACGTCGCGCATACCTGGAAGAGCAATTTGCTGCCCTGA 5236
Db      4190  ACAAAATAACACAGATCCTGACGTCGTACATATTGGAATAAACAAGCCTTCCTCTTC 4249
Qy      5237  GGTTTGGGTCAACATGATCAAGAACCCGCACTTGTGTGATCCATCCATAGAAGACGA 5296
Db      4250  GCTTCTGGGTAAACATCCTGAGAACCTCAGTTGTGTGACATTAAGAAAGACACAC 4309
Qy      5297  TCACAGACGCTGCTCTCTGTGTGAGGCTCAGACCTTCATGAGACTTGTCTCAAGCTAG 5356
Db      4310  ATATGACGGCTGTTTGTTCAGATGATGGCCAGGCACTTATGAGATGATTTCTCTCAAG 4369
Qy      5357  AGCACCGGCTGGGAGAGAGCTGCCCTCCACACAGCTGTGTATGCCAAGACATCCCA 5416
Db      4370  AGCACCACTAGGAGAGAGAGCAACCACTAATTAACCTCTCTATGCCAAGATATCCAA 4429
Qy      5417  GGTACAGAAATGGGTGAGAGATTTACTCAGACATAGGGAAGATGCCACCATCAAGC 5476
Db      4430  CTTACAGAAAGAAATTAATTTATTCAGAAAGCAATCAGGATTTGCTCCTCATGTGAT 4489
Qy      5477  ACCAAGACATGAACCATACCTGGCTGAGCAGTCCCGGATGACATGATAGTTCAACA 5536
Db      4490  CCTCAGAAATGGAAGATTTTAACTCAGGAATCTAGAAACATGAAATGATTTATG 4549
Qy      5537  CCATAGTGCATCTCAGAGATCTTCTCTATGTGGGCAATATACGAGGAGAGATCCT 5594
Db      4550  AAGAACTGCTTGAAGAAATTTACAATACATGTAATAATTTTATGATGATTC 4607

```

```

RESULT 9
US-09-328-475C-37/c
; Sequence 37, Application US/09328475C
; Patent No. 6476207
; GENERAL INFORMATION:
; APPLICANT: Zhang, Jimmy
; APPLICANT: Astel, Jon H.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Endege, Wilson O.
; APPLICANT: Ford, Donna M.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; APPLICANT: Steinhilber, Kathleen E.
; TITLE OF INVENTION: GENES AND GENE EXPRESSION PRODUCTS THAT
; FILE REFERENCE: 1532.002/200130.463
; CURRENT APPLICATION NUMBER: US/09/328,475C
; NUMBER OF SEQ ID NOS: 341
; SOFTWARE: PatsSeq for Windows Version 3.0
; SEQ ID NO 37

```

```

; LENGTH: 1024
; TYPE: DNA
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(1024)
; OTHER INFORMATION: n = A,T,C or G
US-09-328-475C-37
Query Match 1.5%; Score 86.4; DB 4; Length 1024;
Best Local Similarity 67.9%; Pred. No. 2,4e-11;
Matches 148; Conservative 0; Mismatches 68; Indels 2; Gaps 2;
Qy      5472  CAGCGACCAAGACATGAAGACATA-CCTGGCTGAGCAGTCCCGATGACATGAATGAG- 5529
Db      780  CAGTACACAGGCCCATGAATGCTTACCTGCGAGACAGTCCCGCTGCACACCGTGAGT 721
Qy      5530  TTCAACACATAGATGACACTTCAAGATCTTCTCTATGTGGCAATATACGCGAGAG 5589
Db      720  TTCAACATGTTGAGTCCCGCCCAATGATCTACTCTATGTGCAAGTATGTAGAGAG 661
Qy      5590  ATCTTGAACCTCTGAGCACGACATGACACAGTGTGGGAAAGCAAGAACTGGCTTAACTA 5649
Db      660  CTATCGGGGCGCTTACAGAGATGAGACAGGACGCGGCAACGCGCTGCTTATAAGTG 601
Qy      5650  GAACAGTCATTAACCTCATGAGCTTATGACAGCTGAAA 5687
Db      600  GAGCAGCTCATTAATGCAATGTCATGATGAGAGCTGAGA 563

```

```

RESULT 10
US-09-313-294A-6281
; Sequence 6281, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalzudi, Raghnath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 6281
; LENGTH: 288
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. 6476212 700351536H1
; LOCATION: 9, 19, 26, 50, 82, 109, 135, 175-176, 186, 197, 208, 217, 244
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-6281
Query Match 1.0%; Score 57.8; DB 4; Length 288;
Best Local Similarity 73.2%; Pred. No. 0.00013;
Matches 71; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
Qy      4685  GGATGATCTTTCAGAGATGAAGACATACACCAAGATTGAGATGATTTGAGAGCACTGA 4744
Db      12  GGGTCGTCGCAANTGAGAGACATCACCAAGATTNAGGTGACTGGAAAGCGGCTCA 71
Qy      4745  ACACTGCGCCCACTACCAAGTCCAGATGTTCCGT 4781
Db      72  ACACTGATGATTCATTCAGGTGAGAGGGGCTGCTT 108

```

```

RESULT 11
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367

```

GENERAL INFORMATION:
APPLICANT: DOENER, F.
APPLICANT: SCHEITLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
City: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZpT-Fls
US-08-232-463-14

Query Match 0.9%; Score 49.6; DB 1; Length 7218;
Best Local Similarity 4.0%; Pred. No. 0.058; Mismatches 163; Indels 0; Gaps 0;
Matches 16; Conservative 219; Mismatches 163; Indels 0; Gaps 0;

QY 4649 ATATGATCTGATGCGCAAGAGAGTGGGCAAGATGATCTTGACAGATGAAGACA 4708
DB 1443 ATTGGACRRR 1384
QY 4709 TCACCAACAAGATTGAGATTGGAAGAGATGAACACACTGGCCCACTACAGAGTC 4768
DB 1383 RR 1324
QY 4769 CAGATGTCCTGGTGGCATTAAGTTCACAGAGAGTGAAGAGCTTAACGACAGTGAACA 4828
DB 1323 RR 1264
QY 4829 ACTCCACCTCTCCAGACCTCAGCAAGTAATATGAATAATGATCCGTACAGGGCA 4888
DB 1263 RR 1204
QY 4889 GCCCGCAGAGCTCGCTCAGCAGACCTATGATCACTCTCTGAGAGTGAAGTGA 4948
DB 1203 RR 1144
QY 4949 AGATGTGACCTAGTGAAGACAGACGAGACCGAAGAGAGGGGAGACCGGGGA 5008
DB 1143 RR 1084

QY 5009 GCAAGATGTGTCTGAATCTTACTGACCCGACTCTTG 5046
DB 1083 RR 1046

RESULT 12
US-08-607-509-3
Sequence 3, Application US/08607509
Patent No. 5876735
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
TITLE OF INVENTION: METHODS FOR ENHANCEMENT OF PROTECTIVE IMMUNE RESPONSES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
City: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/607,509
FILING DATE: 16-FEB-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MAKI, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.404C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4800
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1867 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 117..1325
US-08-607-509-3

Query Match 0.9%; Score 49.4; DB 2; Length 1867;
Best Local Similarity 45.3%; Pred. No. 0.034; Mismatches 216; Indels 0; Gaps 0;
Matches 179; Conservative 0; Mismatches 216; Indels 0; Gaps 0;

QY 4049 AAGGCTGAACTCTTGCCCAAGCTCATCAACAAGAGTGTCTCTGCTTCAATCC 4108
DB 403 AGGGCTGTGCTCTCCCACTGCGAGCTGGCCCTGCAAGCGAGGTGATCAGCC 462
QY 4109 GCAAGCTTGAATCCAGAGTGTCTTCGATGAGCGGACCGGAGAGCGTGCCTCACTCA 4168
DB 463 GCATGCTGAATCTCTGCAACAGCTTCAGATTCTGCGAGACCTTTGTGGCGGACGCG 522
QY 4169 TCATGACCGTCTGCAAGAGAGTGAAGTGAAGCACTGATGTGCTGAAGCAAGCTGTGG 4228
DB 523 GCGTCAGAGATGACCTGGCAGAGCTGAGCGCGCGTCAATGTCGCTGGGACAGCGCGG 582
QY 4229 CCGAAGCTATGACAAAGAACTGAGAGCAAGAACCAACCCCTAAGCTGTGCTGAGAGCA 4288
DB 583 GCGGCTGTCCGAGCTGATCAAGGTGGCGCTGCGCAGAGTGGCTGGCGGTGCTGG 642
QY 4289 CTGATCAGTGGCTGAGAGAGTGTGACCAATGATGTTACTTCTCTCTACAAAGTTCC 4348
DB 643 TGCTGAGAGAGGCTGATGATGATGCTGTCAAGGCTTTCGCGAGACAGATTTAACAGATCT 702
QY 4349 TCAAGAGTGTGCTGGAGAGCCCTCTTCTCTCTGTTCTGTGCGCATCAAGAGAGAGATG 4408

Mon Feb 23 09:34:02 2004

us-09-964-956-12.rn1

Page 12

Db 703 TCCGCTTCTGCGGAGGATCCAGTGGCGCTTCTCCGCGCAGATGCCGAGAGG 762
Qy 4409 AGAAGGCGCCCATTTAGCCCATCAGCGGAGGCC 4443
Db 763 TACTGAGCTGACGAAGAAGTTCAATGCGCGACCCC 797

RESULT 13

US-08-634-642-3

Sequence 3, Application US/08634642

Patent No. 5879687

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.

TITLE OF INVENTION: METHODS FOR ENHANCEMENT OF

PROTECTIVE IMMUNE RESPONSES

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/634,642

FILING DATE: 18-Apr-1996

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Makl, David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 210121.404C4

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1867 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS

LOCATION: 117..1325

US-08-634-642-3

Query Match 0.9%; Score 49.4; DB 2; Length 1867;
Best Local Similarity 45.3%; Pred. No. 0.034; Indels 0; Gaps 0;
Matches 179; Conservative 0; Mismatches 216;

Qy 4049 AAGGCTTGAAGCTTTGCGCCAGCTCATCAACAAGTGTCTGCTGCTTCTATCC 4108
Db 403 AGGCGCTGCTGCTCTCCCGCACTGCGAGCTGCGCCGTCAGACGCGGAGTATGAGCC 462
Qy 4109 GCAAGCTTGAAGCTTCCAGCTGCTTCCATGCGCGACCGTGGCAACGTGGCTCACTCA 4168
Db 463 GCATCGGTGAGTCTCTTCGAAAGCTCAAGTTCGCGAGACCTTTGTGGCGGACGC 522
Qy 4169 TCATGACCGTCTGTCAGAGCAAGCTGAGTACCGCACTGATGTCTGAAAGCACTCTCG 4228
Db 523 GCGTGCAGAGATGACTGGCGAAGTGCAGGCGCGCGTCACTGTCCTGGCGACCGCG 582
Qy 4229 CCGACCTCATTTGACAAGACCTGAGAGCAAGCAACCTTAAGCTGCTGCTCAAGAGA 4288
Db 583 GCGCGTGTCCGAGCTGATCAAGGCTGCGCGCTGCGCAAGAGTCCCTGCGCTGCG 642
Qy 4289 CTGAGTCAAGTCTGAGAGATGCTGACCAATTGTTTACTTTCTCTCTTCAAGTTCC 4348
Db 643 TGCTGAGAGAGGCTGATGATGCTGTCTGCTGAGGCTTCGCGGACCAAGATTACGATCT 702

Qy 4349 TCAAGAGTGTGTGAGGAGGCCCTTCTCTCCCTGTCTGTGCTCAAGCAGATG 4408
Db 703 TCCGCTTCTGCGGAGGATCCAGTGGCGCTTCTCCGCGCAGATGCCGAGAGG 762
Qy 4409 AGAAGGCGCCCATTTAGCCCATCAGCGGAGGCC 4443
Db 763 TACTGAGCTGACGAAGAAGTTCAATGCGCGACCCC 797

RESULT 14

US-08-989-370-3

Sequence 3, Application US/08989370

Patent No. 6013268

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.

TITLE OF INVENTION: METHODS FOR ENHANCEMENT OF PROTECTIVE IMMUNE RESPONSES

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/989,370

FILING DATE: 12-DEC-1997

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Makl, David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 210121.404C5

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1867 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS

LOCATION: 117..1325

US-08-989-370-3

Query Match 0.9%; Score 49.4; DB 3; Length 1867;
Best Local Similarity 45.3%; Pred. No. 0.034; Indels 0; Gaps 0;
Matches 179; Conservative 0; Mismatches 216;

Qy 4049 AAGGCTTGAAGCTTTGCGCCAGCTCATCAACAAGTGTCTGCTGCTTCTATCC 4108
Db 403 AGGCGCTGCTGCTCTCCCGCACTGCGAGCTGCGCCGTCAGACGCGGAGTATGAGCC 462
Qy 4109 GCAAGCTTGAAGCTTCCAGCTGCTTCCATGCGCGACCGTGGCAACGTGGCTCACTCA 4168
Db 463 GCATCGGTGAGTCTCTTCGAAAGCTCAAGTTCGCGAGACCTTTGTGGCGGACGC 522
Qy 4169 TCATGACCGTCTGTCAGAGCAAGCTGAGTACCGCACTGATGTCTGAAAGCACTCTCG 4228
Db 523 GCGTGCAGAGATGACTGGCGAAGTGCAGGCGCGCGTCACTGTCCTGGCGACCGCG 582
Qy 4229 CCGACCTCATTTGACAAGACCTGAGAGCAAGCAACCTTAAGCTGCTGCTCAAGAGA 4288
Db 583 GCGCGTGTCCGAGCTGATCAAGGCTGCGCGCTGCGCAAGAGTCCCTGCGCTGCG 642
Qy 4289 CTGAGTCAAGTCTGAGAGATGCTGACCAATTGTTTACTTTCTCTCTTCAAGTTCC 4348

Mon Feb 23 09:34:02 2004

us-09-964-956-12.rn1

Page 13

DB 643 TGTGACACGAGGTGATGATGCTGTCTCAGGGCTTCGCGACCAAGATTACGAGATCT 702
QY 4349 TCAGAGAGTGTCTGGGGAGCCCTTCTTCCCTGTTGTGTCATCAAGCAGAGATGG 4408
DB 703 TCCGCTTCTGCGAAGACATCCAGTTCGCTCTTCTCCCGCACGATCCGGAGAGG 762
QY 4409 AGAAGGCCCCATTGACGCCATCACGGCGAGGCC 4443
DB 763 TACTGAGCTGACGAAGAAGTTCAATGCGGACCCC 797

RESULT 15
US-09-398-169-3
; Sequence 3, Application US/09398169
; Patent No. 6660840

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.
TITLE OF INVENTION: METHODS FOR ENHANCEMENT OF PROTECTIVE IMMUNE RESPONSES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/398,169

FILING DATE: 17-Sep-1999
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/989,370

FILING DATE: 12-DEC-1997
ATTORNEY/AGENT INFORMATION:

NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.404CS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1867 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS

LOCATION: 117..1325

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-398-169-3

Query Match 0.9%; Score 49.4; DB 4; Length 1867;
Best Local Similarity 45.3%; Pred. No. 0.034;
Matches 179; Conservative 0; Mismatches 216; Indels 0; Gaps 0;

QY 4049 AAGGCTGAAGCTTTGGCCAGCTCATCAACAAGAGTTCCTCTCTTCAATCC 4108
DB 403 AGGGCTCTGCTCTCCCGCACTGCGAGCTGCGACGCGAGAGTATCAACC 462
QY 4109 GCAGCTGAGTCCGAGGCTTCTTCATCGCGACCGTGGCAAGTGGCTCACTCA 4168
DB 463 GCATCGGTAATCTCTGCGAAGCTTCAAGTTCTGGAAGCTTTGTGGGGGCAAGC 522
QY 4169 TCATGACCGTGTGAGAGCAAGCTGAGTACGCCACTGATGTCTGAAGAGCTGTGG 4228
DB 523 GCGTCAAGAGTACCTGCGAAGCTGCGAGCGCGGCTCATGCTTGGCGTGGCAAGCGCG 582

QY 4229 CCGACTCATTTGACAAAGAACCTGAGAGCAAGAACCACTTAAGCTGTCTCAGAGGA 4288
DB 583 GCCCGTGTCCGACGTGATCAAGCGTGGCCGCTGGCAAGAGTGGCTGGCTGG 642
QY 4289 CTGAGTCAAGGCTGAGAGATGCTGACCAATTGTTTACTTCTCTCTACAAATTC 4348
DB 643 TGTGACGAGGCTGATGATGATCTGTCTCAGGGCTTCGCGACCAAGATTACGAGATCT 702
QY 4349 TCAGAGAGTGTCTGGGAGCCCTCTTCTCTCTGTCATCAAGCAGATGG 4408
DB 703 TCCGCTTCTGCGAAGACATCCAGTTCGCTCTTCTCCCGCACGATCCGAGAGG 762
QY 4409 AGAAGGCCCCATTGACGCCATCACGGCGAGGCC 4443
DB 763 TACTGAGCTGACGAAGAAGTTCAATGCGGACCCC 797

Search completed: February 19, 2004, 23:18:28
Job time : 272 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 20:48:15 ; Search time 1236 Seconds
(without alignments)
16122.863 Million cell updates/sec

Title: US-09-964-956-12

Perfect score: 5691

Sequence: 1 atgaaagccatgcctcgaa.....gcttagacagctgaataaa 5691

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 2308684 seqs, 1750822206 residues

Total number of hits satisfying chosen parameters: 4617368

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEM_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEM_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEM_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEM_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US60_NEM_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_NEM_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	2778.4	48.8	6730 14 US-10-175-523-95	Sequence 95, Appl
2	2533.8	44.5	5895 12 US-10-087-684-91	Sequence 31, Appl
3	2533.8	44.5	5895 12 US-10-218-779-91	Sequence 31, Appl
4	1157.8	20.3	3666 15 US-10-108-260A-802	Sequence 802, Appl
5	793.8	13.9	2597 14 US-10-245-103-91	Sequence 91, Appl
6	793.8	13.9	2597 14 US-10-245-107-91	Sequence 91, Appl
7	793.8	13.9	2597 14 US-10-245-143-91	Sequence 91, Appl
8	793.8	13.9	2597 14 US-10-245-771-91	Sequence 91, Appl
9	793.8	13.9	2597 14 US-10-245-851-91	Sequence 91, Appl
10	793.8	13.9	2597 14 US-10-245-883-91	Sequence 91, Appl
11	793.8	13.9	2597 14 US-10-237-535-91	Sequence 91, Appl
12	793.8	13.9	2597 14 US-10-238-183-91	Sequence 91, Appl
13	793.8	13.9	2597 14 US-10-238-283-91	Sequence 91, Appl
14	793.8	13.9	2597 14 US-10-238-370-91	Sequence 91, Appl
15	793.8	13.9	2597 14 US-10-245-055-91	Sequence 91, Appl

16	793.8	13.9	2597 14 US-10-245-147-91	Sequence 91, Appl
17	793.8	13.9	2597 14 US-10-245-730-91	Sequence 91, Appl
18	793.8	13.9	2597 14 US-10-245-739-91	Sequence 91, Appl
19	793.8	13.9	2597 14 US-10-246-210-91	Sequence 91, Appl
20	793.8	13.9	2597 14 US-10-239-196-91	Sequence 91, Appl
21	793.8	13.9	2597 14 US-10-243-024-91	Sequence 91, Appl
22	793.8	13.9	2597 14 US-10-243-409-91	Sequence 91, Appl
23	793.8	13.9	2597 14 US-10-245-621-91	Sequence 91, Appl
24	793.8	13.9	2597 14 US-10-245-880-91	Sequence 91, Appl
25	793.8	13.9	2597 14 US-10-245-033-91	Sequence 91, Appl
26	793.8	13.9	2597 14 US-10-243-095-91	Sequence 91, Appl
27	793.8	13.9	2597 14 US-10-245-185-91	Sequence 91, Appl
28	793.8	13.9	2597 14 US-10-245-427-91	Sequence 91, Appl
29	793.8	13.9	2597 14 US-10-245-473-91	Sequence 91, Appl
30	793.8	13.9	2597 14 US-10-245-770-91	Sequence 91, Appl
31	793.8	13.9	2597 14 US-10-245-877-91	Sequence 91, Appl
32	793.8	13.9	2597 14 US-10-246-876-91	Sequence 91, Appl
33	793.8	13.9	2597 14 US-10-243-320-91	Sequence 91, Appl
34	793.8	13.9	2597 14 US-10-242-743-91	Sequence 91, Appl
35	793.8	13.9	2597 14 US-10-242-845-91	Sequence 91, Appl
36	793.8	13.9	2597 14 US-10-237-636-91	Sequence 91, Appl
37	793.8	13.9	2597 14 US-10-238-325-91	Sequence 91, Appl
38	793.8	13.9	2597 14 US-10-238-346-91	Sequence 91, Appl
39	793.8	13.9	2597 14 US-10-243-411-91	Sequence 91, Appl
40	793.8	13.9	2597 14 US-10-243-424-91	Sequence 91, Appl
41	793.8	13.9	2597 14 US-10-243-445-91	Sequence 91, Appl
42	793.8	13.9	2597 14 US-10-243-446-91	Sequence 91, Appl
43	793.8	13.9	2597 14 US-10-245-874-91	Sequence 91, Appl
44	793.8	13.9	2597 14 US-10-242-653-91	Sequence 91, Appl
45	793.8	13.9	2597 14 US-10-243-167-91	Sequence 91, Appl

ALIGNMENTS

RESULT 1
US-10-175-523-95
Sequence 95, Application US/10175523
Publication No. US20030096264A1
GENERAL INFORMATION:
APPLICANT: Brockman, Jeffrey
APPLICANT: Evans, David
APPLICANT: Hook, Derek
APPLICANT: Klimczak, Leszek
APPLICANT: Laeng, Pascal
APPLICANT: Palfreyman, Michael
TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)
FILE REFERENCE: 3235/11795-US3
CURRENT APPLICATION NUMBER: US/10/175,523
CURRENT FILING DATE: 2002-06-18
PRIOR APPLICATION NUMBER: US 60/299,151
PRIOR FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: US 60/317,828
PRIOR FILING DATE: 2001-09-07
PRIOR APPLICATION NUMBER: US 60/325,150
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US 60/333,047
PRIOR FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: US 60/349,936
PRIOR FILING DATE: 2002-01-18
PRIOR APPLICATION NUMBER: US 60/361,834
PRIOR FILING DATE: 2002-03-04
NUMBER OF SEQ ID NOS: 197
SOFTWARE: PatentIn version 3.1
SEQ ID NO 95
LENGTH: 6730
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)-(6730)
OTHER INFORMATION: where n may be a or g or c or t/u, unknown, or other

US-10-175-523-95

Query Match 48.8%; Score 2778.4; DB 14; Length 6730;
Best Local Similarity 69.3%; Pred. No. 0;
Matches 3858; Conservative 0; Mismatches 1691; Indels 15; Gaps 5;

QY 137 AGCCCGCCGAGGTTTCATACCTGCTGTGTGATGAGAGACAGACACATTACTTGG 196
DB 683 ATGTGACTGACCTTTCAACATTGACTGTACCCAGACAGAGAGGGGCTGTGTATGTGG 742
QY 197 GGGCCCTGATCGGATTTTACAGCTCTCCAGCAGCTGGAAGCTTGTGTATCCGATGAGA 256
DB 743 GGGGTATCATGCTGTCTACAAAGTTGACTGGCAACCTCACCTCAGGTGGCTCAAGAG 802
QY 257 CAGGGCCGAGAGAGACACCCCAAGTTTACCACCCCGCATGCTCAGACCTGCAATG 316
DB 803 CAGGGCCGAGAGAGACACAGAGGCTTGTACCCACCTCATTTGTACAGCCTGTGAGTG 862
QY 317 AGCCCTGACCAACCAACATGTCAACAGATGCTCTCTATAGATTAAGAGACCA 376
DB 863 AAGTCTTACACTACCAACATGTCAACAACTACTGATCATTTGACTCTGAGATC 922
QY 377 GAGTGAATGCTGTGGAGCTGTACCAAGGATCTGCAAGCTGTGAGGCTGGAGACC 436
DB 923 GCTGTGCTGTGGAGGAGCTTACCAAGGCTGTTCAGAGCTCTGCGACTGATGATGACC 982
QY 437 TCTTCAAGCTGGGGAGCTTATCATTAAGAGACATCTGTCTAGGTGTCAAGCA 496
DB 983 TCTTCACTGTGTGAGGACATCCCAAGAGAGACATCTTGTCCAGGTCAATAGGA 1042
QY 497 GGGGCTAGCTTTGAGTATGCTCTCTACAGCAACCTGAGATGACAGCTGTCTATG 556
DB 1043 CAGGACCATATATGT 1102
QY 557 CCACGGCAGTGAATGGAGCCCGAGTATTTTCCACATCTCCAGCCGAGAACTGACCA 616
DB 1103 GCATGCTGTGATGGAGAGGAGATTACTTCCCTACTGTGTCCAGCCGAGAGTGTGCC 1162
QY 617 AGAAGCTGAGGCGAGATGCGATGTCGGTACGCTTCGATGATGATGATGATGATGAT 676
DB 1163 GTGACCTGATCTTCAAGATGTGTGATGTGATGTGATGTGATGTGATGTGATGTG 1222
QY 677 TGATTAAGATCCCTCGAGACCTTCAACATCATCCCTGACTTGTATCTACTATGTCT 736
DB 1223 TCATCAAGATTCCTCTGACACCTGAGCCCTGTCTCTCACTTGCATCTTCAATCTCT 1282
QY 737 ATGTTTATGAGTGGCACTTGTCTACTTTTGAACCTCCAGACCTGAGATGTGTCTC 796
DB 1283 ATGGCTTTCAGATGGGGGTTGTCTACTTCTCACTGTCTCAAGATGTGTGTCTGCA 1342
QY 797 CACCAAGCTTCA---CCACCAAGAGAGAGTGTATATCATCCAGCTGTGTAGGCTTTGCA 853
DB 1343 GCAATGGCATCAATTCACCTGAGAGACCTTCTATACTCAAGATGTGTGTCTGCA 1402
QY 854 AGGAGAGACAGGCTTCACTCTCTATGTAGAGTGGCCATGGCTGTGAGAGGAGAGGG 913
DB 1403 AGGATGAGCCCAAGTTCACTCTCTATGTGTCTCTGCTTTGTGGCTGCAACAGTGTCT 969
QY 914 TGAGATACCGGCTGTGAGGCTGTCTACTGTCCAAAGCCGGGGCGTGTGGAGAGA 973
DB 1463 TGGATATATGCTTGTGAGAGAGTCTTACTTGTCAAGAGCCAGGGAGAGCTTACTGAG 1522
QY 974 CCTTGAAGTCAATCAAGTATGATGATGCTGTCTTACCGTCTTCTCCAAAGGCGCAAGC 1033
DB 1523 CCTTCAACATCAGACGAGCAAGATGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1582
QY 1034 GGAATATGAATCCCTGATGATGCTGGCCCTGTGATCTTCACTTGTGAAGAGATTAATG 1093
DB 1583 AGTACACAGACCCCTGTATGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1642
QY 1094 ACCGCAATTAAGAGAGGCTGTGATGCTTGTATCCGGGGGAGAGGAGAGCTGTGAGCT 1153
DB 1643 TGCATATCAAGAGAGGCTGTGATGCTGTGTATCAAGAGAGGCAACTTGTGAGCTCACT 1702

QY 1154 GGTCAAGTGAAGAGACATCCCTGAGCAGAGTGTGCTCTTAAACATTGACATTAATCTT 1213
DB 1703 GGTGTGTGGAAAGAGATGTGACAGTGCACAAAGGCGCTGTGCCAATGATATTAATCTT 1762
QY 1214 GTGGCTTGAATGAATATCTCCCTGGAGAGTGTCCGACATGTGTGTGTGTGTGTGTGT 1273
DB 1763 GGGGCTTGAATCAACAGCTTGTGGAGGCTCACTCTGTGTGTGTGTGTGTGTGTGTGTGT 1822
QY 1274 TCACGAGAGACAGGAGCCGATGACATGTCTGATGATATGTCTTAAAGAACATCTC 1333
DB 1823 ATACACAGAGAGAGAGCCGCTGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1882
QY 1334 TGGCTTTGTGGGACCAAAAGTGGCAAGCTGAAAGATTCGGGTGTATGAGACCAAGG 1393
DB 1883 TGGTTTTGTGGGAGCTAAGAGTGGCAAGCTGAAAGATTCGATGTGTGTGTGTGTGTGT 1942
QY 1394 GCAAGCCCTCAGATGAGACGAGTGCAGTGT---GGAACCCCGCCCACTGTCCGGG 1450
DB 1943 ATGTGTGGGTCCAGTATGATGATGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2002
QY 1451 ATATGCTTGTCCAGAGACACAGACACTCTATATGTCTGAGAGAGAGCTCAACCA 1510
DB 2003 ACATGCTTCTTCCATCAATGATGATATCTTATGTATGTATGTGTGTGTGTGTGTGT 2062
QY 1511 GAGTCCCTGTGGAGTCTGT 1570
DB 2063 GGTGTCTGT 2122
QY 1571 ACCCCACTGT 1630
DB 2123 ATCTCACTGT 2182
QY 1631 GGTCAAGAGAGCCCGAGGTTTGGCTCGAGATGAGACAGTGTGTGTGTGTGTGTGTGT 1690
DB 2183 GGGCTGTGGAGAAATGATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2242
QY 1691 ATCCCAACATATCTCGTCTCTCAAGTAAAGAGTGTGTGTGTGTGTGTGTGTGTGTGT 1750
DB 2243 ACCCCCAACAGATCTGT 2302
QY 1751 TCCCGAGTGTCACTGT 1810
DB 2303 CTCCCAACTCTGT 2362
QY 1811 TGT 1870
DB 2363 AGGTATGT 2419
QY 1871 TCACAGAGATGGAGACCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1930
DB 2420 TCCCTGT 2479
QY 1931 TGACCTTCCGAGACACAGCTTGTCTTCAATTTGACAGGTGTGTGTGTGTGTGTGTGTGT 1990
DB 2480 AGATCTTGT 2539
QY 1991 CCTGT 2050
DB 2540 CCTGT 2599
QY 2051 ACCCAAGAGCTGTCTCTTCCAGAGAGCCGAGTGAAGTGTGTGTGTGTGTGTGTGTGTGT 2110
DB 2600 ACCCACTACTGTCTCTTCCAGAGAGAGAGATCAATGTGTGTGTGTGTGTGTGTGTGTGT 2659
QY 2111 TGT 2170
DB 2660 TGT 2719
QY 2171 CCAAGAGCTTCCCAAGCCCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2230
DB 2720 CCGAAGCTTGT 2779

QY	2231	AGGGCAGCGAGCAGCAGAGTGGCCCGCCCGGCTTCAACAGCTCCAGCGTAAGTGGCCGA	2230
Db	2780	AAGGGGCTGTCCACCGGGTCCCTGGCTTCGCTTCAACAGTTCAGTGTGACGTGCCAAA	2839
QY	2291	ACACCTCTTATTCATAGAGGAGATGGAGATCAACACCTCCCGTGGAGTTGACAGTGC	2350
Db	2840	ACAGCTGTACCAAGATATGATGGAGATGAGCATTCAGAACCTTACAGATGAGCTTTGGTAG	2899
QY	2351	TGTGGAAATGGGCACTTCAACATTTGACAAACCAGCTCAGAAATTAAGTTTACCTTTACAGT	2410
Db	2900	TATGGAAATGGCACTTCAATTAATGCAACCTCAGAGACTGAAGTATATCTTACAAAGT	2959
QY	2411	GTGAGCCATGCGTAGAGACTGCGGGGCTGTGCTCAAGCTGACCCAGCTATCCGATATGTG	2470
Db	2960	GTGACGCCCAAGGGGAAAGCTGTGTGTCTCTGCTCAAGGCTGACCAAGTTTCAAGTGTG	3019
QY	2471	GCTGTGTCCAGGGCCCAAGGCCAGATGTCACCTTGCCTCAGACACTGCTTGGCCAGAGAGCC	2530
Db	3020	GCTGTGTGACGTGTGTAGCGCAGATGTATCCCTTCCACAGACACTGCTCCAGCACTTGTAGCC	3079
QY	2531	AGTGGCTGAGAGCTGTCTGTGTGCAAAAGCAAGTGTGACAAACCCCGGCACTACAGAGATAA	2590
Db	3080	CCTGGCTTGACTGTGTCCAGCCAAATGTCAAAAGTGTCCAAACCCCAATCCAGAGATTT	3139
QY	2591	TCCCGGTGACAGGCCCCCGGAGAGGGGACCAAGTGCATATCCGAGGGAGAACCTGG	2650
Db	3140	TGACAGTATCAGAGAACCACTGAGAGGAGGAGCTCGGTATCACCATTCATGGCTGAACTGG	3199
QY	2651	GCTTGGAAATTTCCGACATCGCCCTCCCATGTTCAGAGTGTCTGGGGTGGAGTGCAGCCCTT	2710
Db	3200	GCTTGGACTTCTCTGAGATGTGTCAACCATGTGCAAGTGTGCTGGAGTGGCTCTTGACACTTA	3259
QY	2711	TAGTGAATGTATTCATCCCTCGCAGAACAGATCGTGTGTGAGATGGGGAGGCGCAAGCCCA	2770
Db	3260	TCCCAAGGGGAAATACATCATCGCTGAGCAGATGTGTGTGAGATGGGCCCATGCCATTATAG	3319
QY	2771	GCCAGCATGCAAGCTTGGTGTGAGATCTCGTGGCTGTGTGTGCGCTTGAATTCATGGCCC	2830
Db	3320	GTACCACATCTGGGGCTGTGGGCTGTGTCATTTGGGGAAATGGAAAGCAAGTTCATATGACCA	3379
QY	2831	GATCCTCAGAGTCATTAATCTTCAATGACACTGACCTCTCAGATGTGAAGCCCAAGCCGGG	2890
Db	3380	AGTCCCAACAGCAGATACCTTTTGTGAATCTCTGTGTCTGTCACTCAGCCCGATCCGGG	3439
QY	2891	GAGCCATGTCCGAGAGGAGCCCAAGTACATTCACAGGCAACCAACTGAATCCCGAAGCA	2950
Db	3440	GACCAGAGTCAAGGAGGTACCAATGTGGACCAATTCACAGGCCATTTACTTGTCTGGAGCA	3499
QY	2951	AAGTGTGTGTATGTTTGGAAAGCAGCCCTGTCTTCCACAGGGAGATCTGCATCTTCCA	3010
Db	3500	GTGTGTGCAGTGTACTGTGGCAATCAGACCTGTGAATTTTAATGGAGGTCAATGAATGAGA	3559
QY	3011	TTGTGTGCACACA--CCACATCTTCAGATGAGTGTCTAGATGAATGAAGTGTGGTGCAGG	3067
Db	3560	TTGTATGTGTTCACCCCATCATCATCATATGACGTAGACCAAGTCCCTGTCTCCGTGAGTGTG	3619
QY	3068	TGACACAGGCCCAAGATTCACACAGAGACCTGTGTCTTCAATATGTGAAACACCCCAACTG	3127
Db	3620	TGCAACAGGCCCGGGTGTGATGACAGTCTGCAGTTGTGAATTAATATATACCCAGGGTCC	3679
QY	3128	TGGCGATTGAGCACAATAGAGACATTTGTACGTGAAACACACCCATCCCGTATGGGGGA	3187
Db	3680	AACGTATTTAGCCCAAGTGTGAGTATCACATGTGTGGCAACACCCCTTAACATCACAGGCT	3739
QY	3188	CCACACTGCACTCATACAGAACCCCAAGATCTGTGCAAGCATGTGAAGGAAGAGACCA	3247
Db	3740	TCAACTTGTGATGTCAATTCAGAGGCCCAAGGTTCCAGTCAAAATTTAATGGCAAAAGATCTG	3799
QY	3248	TCAATATCTGTGAGGTTCTGAAACGTACTGAGATATACCTGTACAGGCGCCCGCCCTCGCTC	3307
Db	3800	TCAATGTATGCACAAGTGTAAACAAACACCAACCTTCACTGTGTGGCAACCTCTCTGTACCA	3855
QY	3308	TGGGTCTGACCAACAGTCAAGCTGACCCGAGAGGCCGAGAGATTGTGACTTATCTGTGG	3367

Db	3660	GTGACTACGGCCACAGCTCTGGACACTGTGGAAACGGCCAGATGAGTTTGGATTCTCTTTA	391.9
Ox	3368	ACAACTCCAGATCCCTGCTCATCTCAACAAGACAACTTCACTTAATTCACACCCGG	342.7
Db	3920	ACAACTGTTCAATCTTACTCATCTATACACACACAAAGTTTCATCTACTACCCCAACCA	397.9
Ox	3428	TGTTTGAAGCCTTTGATCCCTCAGAGATCCCGAGACTCAAGCCCTGGACGCCCATCATCC	348.7
Db	3890	CGTTTGAATCTGCTACGCCCACTGGATCTTGGATCAGAAACCAAGGCTCAACCATCATCC	403.9
Ox	3488	TAAAGGCCAAGAACCTTGATCCCGCTGTGGCTGGGGGCAACGTGAAGCTGAATCACTG	354.7
Db	4040	TGAAGGCCAAATAATCTCTGTCTCTCTCTCTGGAG---GGCCAAACTCACTACACAG	409.6
Ox	3548	TGCTGTGTGGGAGAAACCCCTGACCGGTGACCGGTGACATGTCACATGCTGTCCAGAT	360.7
Db	4097	TAAATATTGGAGAACACTTGTACATCTGTGTGAGAACACACACTGCTTTGTGAC	415.6
Ox	3608	CCCCCAACTCATGGCAGGACCAAAATGATGGCCCTGTCTGGTGAGATGAATCTCC	366.7
Db	4157	CTCCCAACCTCACAGGACAGACAAAGTCAATGTTCACTGAGGGAGGATGGTCTTCAC	421.6
Ox	3668	CGGGAGATGTGTACTTGGCCCGGACAGCCCGCTCAAGCTTCCCGGCATGTCAAGCATCG	372.7
Db	4217	CTGGCTCCGTGAGGTCAATCTCCGACAGCTGTGTGACCTTCGACGCATCATCCAGATCG	427.6
Ox	3728	CAGTGGCTGGCGGCTCCTCATCTTTTCAATCGAGCCGTGCTATTGCTATTAAGCA	378.7
Db	4277	CAGCTGGTGAAGCTCTCTTCTATCATCTGTCACTATGTCATCATGTTCTCATGCTTACAGGCA	433.6
Ox	3788	AGTCCCGGAAAGTACCTTACGCTGAACGGCTGCAGATGCAGATGGAACACTGGAAT	384.7
Db	4337	AGTTAAGGAAATGACCTCACTCAACGGCGCTCCAAATGCAATGGAACAACCTGAGT	439.6
Ox	3848	CCCGTGTGGCCCTGGAATGCAAGGAACCTTTGCCAGCTGCAGACGAGACATCATGAGC	390.7
Db	4397	CCAGGGTGACCTGAGATGACAGGAACCTTTTGGAGCTTCAGACAGACATCATATAGC	445.6
Ox	3908	TGACCAATGACTTGATGAGACCGGAAATTCGTTCTCGATCTGAATAGAACTTACACATGC	396.7
Db	4457	TAAACAATGACTTGATGACTGATACAGGAATCCCTTACTGGAATACCGTATCCATGACATGA	451.6
Ox	3968	GGGTGCTGTTCCAGGAATTGAAGACCAACCTGTCTCCGGACCTTGAAGTCCCGGGCT	402.7
Db	4517	GAGTCTGTTCCAGGCAATGAGGACCACTGTTCTCGCGGAACTGGAAGTACAGGAA	457.6
Ox	4028	ACCGGCAGAGCGGTGTGAGAAAGGCTTGAAGCTTTCGCCCAGCTCATCAACAACAGG	408.7
Db	4577	ATGACACAGACAGATGTGAGAAAGCCCTGAAACTTTCGCCACGCTTATCAACAACAAG	463.6
Ox	4088	TGTTCTGCTGTCTCTTATCTCGACACGTTGAGTCCAGCTGATGTTTCCATGGGCGAC	414.7
Db	4637	TGTTCTTCTGACCTTATCCGTACACTGGAACTACAGCGGAGGTTTCTCATGCGGAAC	469.6
Ox	4148	GTGGCAACGTGGCTCACTCATATGACCGGTGCGACAGCAAGCTGAGTACGCCCATG	420.7
Db	4697	GTGGGAACGTGGCTCTCTATATATGACAGGCCCTTACAGGCTCGGCTAGAAATATGCCACTG	475.6
Ox	4208	ATGTGCTGAAGACGCTGCTGGCCGACTCATTTGACAAGAACCTGGAAGACAAACCAAC	426.7
Db	4757	ATGTCTCAAGACAGCTCTCTCTACCTCAATTGCAAGAACTGTGAGAAACAAGACAC	481.6
Ox	4268	CTAAGCTGCTGCTCAGGAGGACTAGTCAAGGTGTGAGAAAGATCTGACCAATGTGTTA	432.7
Db	4817	CCAAGCTGCTTCTCCGAGACTAGTCTGTGGCCGAGAAAGATCTGACTATACGTGTTG	487.6
Ox	4328	CTTTCCTCTCTACAAATCTCTCAAGAGATGTGTGCGGAGCCCTCTCTTCCCTGTCT	438.7
Db	4877	CTTTTCTTTCAACAATAATCTCTGAAAGATGTGTGTGGGAAACAACCTCTTCACTGTATATCT	493.6
Ox	4388	GTGCATTCAGACAGATGGAAGAGGCCCATTTGACGCATACGGGACAGGCCGCT	444.7

4937 GTGCATCAAGCAGCAGATG3AAAAAGGCCCATTTGACGCTATTACTGTGTAGGCCCAT 4996
4448 ACTTCCTTGAGCGAGCAAGCTTCATCGGCAGAGATTGACTACAAAACCTGTCTCTGA 4507
4997 ACTTCCTTGAGCGAGCAAGCTTCATCGGCAGAGATTGACTACAAAACCTGTCTCTGA 5056
4508 GCTGTGTCAAGCCGAGCAATGCCAAGCCGAGGCTCCAGTAAAGTCTCAACTGTG 4567
5057 ACTGTGTCAAGCCGAGCAATGCCAAGCCGAGGCTCCAGTAAAGTCTCAACTGTG 5116
4568 ACACATCATCTAGCTCAAGAGAGATTTGTGATGCTTCAAGAAATGTCCTGTGCT 4627
5117 ACACATCATCTAGCTCAAGAGAGATTTGTGATGCTTCAAGAAATGTCCTGTGCT 5176
4628 CCACCGGCGCCAAAGCTGAGATATGATCTGAGTGGCGACAAGGAAGTGGGGCAAGGA 4687
5177 CCACCGGCGCCAAAGGCTGAGATATGATCTGAGTGGCGACAAGGAAGTGGGGCAAGGA 5236
4688 TGATCTTGAGAGATGAAGATATCAACCAAGATTGAGATGATTTGAGAGCGACTGAA 4747
5237 TGGGTGTGAGAGAGAGATTAACCAAAATGAGGTGACTGAAAGCGGCTTTAACA 5296
4748 CACTGGCCCACTACAGAGTCCAGATGTTCCGTGTGAGTATGTTCCAGAGAGTGA 4807
5297 CACTGATGATTAACAGAGTGTCAAGAGATCCGTGTGAGTATGTTCCAGAGAGTGA 5356
4808 CAGCTTAAAGCAGATGAACAATCCACCGCTTCCAGAGACTTCAAGAAATATGAAA 4867
5357 CACTCTCAACATCTCCTGCTGCTCCAGCATCTCTGGAATCATTAAGAGTATGACT 5416
4868 ACATGATCCGGGTACACGGGAGCCCGCAGACCTCCGCTCAGCAGACCTATGATCATC 4927
5417 CTTCTCTCAAGTATACAGAGAGCCCGCAGACCTCCGCTCAGCAGACCTATGATCATC 5476
4928 CTGACCTGAGAGATGAGTCAAGTGTGCACTTATGAGAAACACAGAGCAGAGAGACC 4987
5477 CAGACTTGAGAGAGGAGTCAAGGTTTGGCATCTGTGAGAAATCATTAAGAGTATGACT 5536
4988 AGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5047
5537 AGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5596
5048 CCATTAAGGAGCAGCTGAGAGATTTGTGATGACTCTTTGAGACCATTTTGAACAAG 5107
5597 CCACCAAGGAGCAGCTGAGAGATTTGTGATGACTCTTTGAGACCATTTTGAACAAG 5556
5108 CACACCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5167
5657 TGCACCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5716
5168 AGGCTGATTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5227
5717 AGGAGAGACAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5776
5228 TGCCCTTGAAGTTTGGGTCAACATGATCAAGAACCGGAGTTTGTGTTGACATCCATA 5287
5777 TTCCATCTGTTTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5836
5288 AGAAGAGATACAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5347
5837 AGGAGAGATACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5896
5348 CCACCTGAGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5407
5897 CCACATCAAGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5956
5408 ACATCCCAAGTACAGAAATTTGGGTGAGAGGATTTACTCAGACATAGGAAGATGCTCAG 5467
5957 ATATCCCAAGTATTAAGAACTGGGTAGAAATGATGATGATGATGATGATGATGATGATG 6016
5468 CCATCAAGCAGCAAGATGAACGATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5527
6017 CCATTAAGTACCAAGATATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6076

5528 AGTTCAACACCATGAGTGCATCTCTCAGAGATTTCTTCTTATTTGGGCAATACAGGAG 5587
6077 AGTTCAATATGCTGAGGCGCCCTCAACGAGATCTTACTATATGTCAGCAAGTACAGTAGG 6136
5588 AGATCCTTGAGCCTCTGAGCCAGATGACAGATGAGTGGGAGAGCAAACTGGCTTACAAAC 5647
6137 AGCTTATCGGAGCACTGAGAGAGATGAACAGGCGGAGCAGCAGCTGCTTAAAGG 6196
5648 TAGAACAAGTATPACCTCTATGAGCTTAAACAGCTGAATTA 5691
6197 TGAGCATCTCATCAAGCCATGTCATAGAGCTGAAAGGA 6240

RESULT 2
US-10-087-684-31
Sequence 31, Application US/10087684
Publication No. US20040029116A1
GENERAL INFORMATION:
APPLICANT: Edinger, Shlomit R.
APPLICANT: MacDougall, John R.
APPLICANT: Miller, Isabelle
APPLICANT: Ellerman, Karen
APPLICANT: Stone, David J.
APPLICANT: Grose, William M.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
APPLICANT: Caeman, Stacie, J.
APPLICANT: Spytek, Kimberly A.
APPLICANT: Boldog, Ferenc L.
APPLICANT: Li, Li
APPLICANT: Padigaru, Muralidhara
APPLICANT: Mishra, Vishnu
APPLICANT: Shenoy, Suresh G.
APPLICANT: Rastelli, Luca
APPLICANT: Tchernev, Velizar T.
APPLICANT: Vernet, Corine A.M.
APPLICANT: Zernusen, Bryan D.
APPLICANT: Malpankar, Uriel M.
APPLICANT: Guo, Xiaojia
APPLICANT: Miller, Charles E.
APPLICANT: Gangoli, Esma A.
TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-214 CIP
CURRENT APPLICATION NUMBER: US/10/087,684
CURRENT FILING DATE: 2003-03-10
PRIOR APPLICATION NUMBER: 60/253,834
PRIOR FILING DATE: 2000-11-29
PRIOR APPLICATION NUMBER: 60/250,926
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: 60/264,180
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/274,194
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/313,656
PRIOR FILING DATE: 2001-08-20
PRIOR APPLICATION NUMBER: 60/327,456
PRIOR FILING DATE: 2001-10-05
NUMBER OF SEQ ID NOS: 220
SOFTWARE: CnraSeqList version 0.1
SEQ ID NO 31
LENGTH: 5895
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (23)..(5797)
US-10-087-684-31
Query Match 44.5%; Score 2533.8; DB 12; Length 5895;
Best Local Similarity 68.0%; Pred. No. 0;
Matches 3812; Conservative 0; Mismatches 1712; Indels 81; Gaps 17;

QY 141 CGCGAGGGTTTCATACCTGGTGTGATGAGAGACAGACACATTACTTGGGGGC 200
Db 217 CGACTGGGGCTCCACCCACTAATGTGATGAGACAGAGGAGGTGATATGTGGCCC 276
QY 201 CGTCAATCGATTTTACAAGCTCTCCAGCACTGGAAGTCTTGGTACCGATAGACAG 260
Db 277 AGTGAACCGCATCTATAGCTGTGTGGGAACTGACACTGTCCGGGCCACGTCAAG 336
QY 261 GCCGAGAGAGACAAACCCCAAGTGTACCCACCCCGATGTGTCAAGCTGTCAATAGCC 320
Db 337 CCTGTGAGAGACAAACAGAGTGTACCCCGCCCGACGTGTCAATCTGTCCCCACAG 396
QY 321 CTTGACCAACCAACCAATGTCAACAGATGTCTCTATAGACTTACAGAGAAACAGCT 380
Db 397 CTTGGGCAATGACCAACAGTCAACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 456
QY 381 GATTGCTGTGGGGCTGTACCAAGGATCTGTCAAGCTGTGTGTGTGTGTGTGTGTGT 440
Db 457 GCTGGCTGTGGAGCGCTCTCCAGGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 516
QY 441 CAAGCTGGGGAGCTTATCTATAGAAAGACATCTGTCAAGTGTCAACGAGCGG 500
Db 517 CAATCTGGGTGAGCAACCAACGTAAGGACACTTCTGTCAAGCTGTGTGTGTGTGTGT 576
QY 501 CTCAGTCTTTGAGT 554
Db 577 CAGCATGGCGGGCTGTCTCATTTGCGGGCCACCGGGCCAGGGCCAGGCTTTTCTGT 636
QY 555 TGCCACGGCAGTGTGAGGAAAGCCGAGTATTTTCCACATCTTCAGCCGGAACCTGAC 614
Db 637 GGGACACCCATGATGGCAAGTCCGAGTACTTCCCACTGTGTGTGTGTGTGTGTGTGT 696
QY 615 CAAGATCTGTGAGCGGATGTGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 674
Db 697 GGCACACGAGAGATGCCACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 756
QY 675 GATGATTAAGTCCCTTCGACACTTCAACATCATCTGTGACTTGTATCTACTATGT 724
Db 757 ACAGTCAAGATCCCTTCGACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 816
QY 735 CTATGTTTGTGAGTGTGCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 794
Db 817 GTACAGCTTCCGACGCGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 876
QY 795 TCCACAGGCTCCACACAGAGAGAGGTGTATATCAAGCTGTGTGTGTGTGTGTGT 854
Db 877 GACCTGCTGT 936
QY 855 GAGAGACACAGCTTCACTCTATATGAGGTGCCATTGGCTGTGAGCGCAATGGGCT 914
Db 937 GACACACCCCAATCTCTATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 996
QY 915 GAGATACCGCTGTGTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 974
Db 997 GAGATACCGCTGTGTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1056
QY 975 CTTTGGATTCATCCAGATGTATGACTGTCTTCAACGCTTCTCAAGGGCCAGAAAGCG 1034
Db 1057 GCTGGGCTGT 1116
QY 1035 GAAATATGAATCCCTGT 1094
Db 1117 CGCGCTGAGCCACCAAGAGGTCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1176
QY 1095 CGCATTAAGAGCGGCTGTGAGTCTTGTACCGGGCCAGGGCAAGCTGTGAGCTGTGCTG 1154
Db 1177 GAAATTAAGAGCGGCTGTGAGTCTTGTACCGGTGTGTGTGTGTGTGTGTGTGTGTGT 1236
QY 1155 GCTCAAGGTGAAGACATCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1214
Db 1237 GCTGTCTCAACAGAGGCTGT 1296

QY 1215 TGGCTGGACATGAATGCTCCCTGGGAGTGTCCGACATGTGTGTGTGTGTGTGTGT 1274
Db 1297 CGGGCAGGACTTTCAACGAGCTCTGGGGGACAGTCAACATTAAGGAGAGCGCTGTGT 1356
QY 1275 CACGAGGACAGGGAGCGCATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1334
Db 1357 GGTGAACAAGATGATGT 1416
QY 1335 GGCCTTGTGGGCACCAAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1390
Db 1417 GGTATTCGCGGACAGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1476
QY 1391 -----GGGCAACGCTCCAGTATGAGACGTTGAGGTGTGTGTGTGTGTGTGTGT 1445
Db 1477 CGGTGGCGGCTGT 1536
QY 1446 CGGGATATGGCTTTCTTCAAGACCAAGACCACTTATCATGTGTCAAGAGCGAGCT 1505
Db 1537 GCGAATCTGT 1596
QY 1506 CACCAAGTCTGT 1565
Db 1597 GACGGGT 1656
QY 1566 AGGCAACCCCACTGT 1625
Db 1657 AGGCAACCCCACTGT 1716
QY 1626 TGAAGGTGTCAAGAGAGCGCCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1685
Db 1717 TGAAGGACAGAGAGCGCCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1776
QY 1686 GGTTCATCCCAATATCTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1742
Db 1777 TGTGAGCCCGCAATGT 1836
QY 1743 GTACATGTCCCGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1802
Db 1837 CTGGAAGT 1896
QY 1803 GGAATGGGT 1862
Db 1897 TGAAGAGT 1956
QY 1863 CGGATCA-----TCAAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1916
Db 1957 GCCCATCAAGCGGGGACAGGGTGAAGAGACAGGGGTGTGTGTGTGTGTGTGTGTGT 2016
QY 1917 AAGAGAGCCGCAATGACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1976
Db 2017 CAAGAGAGAGAGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2076
QY 1977 CAATTG---TGT 2033
Db 2077 CAGTGTAGT 2136
QY 2034 GATGT 2093
Db 2137 CCAAGT 2196
QY 2094 CGAGAGT 2153
Db 2197 TGAAGATGT 2256
QY 2154 GCTTATCAAGCTGT 2213
Db 2257 ACCCATCAAGCTGT 2316
QY 2214 ATGCATCTTCAATTCAGGGCAGAGAGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2273
Db 2317 GTGTCTTGT 2376
QY 2274 CAGGT 2333

Db	2377	CAAGCTCAGAGCCAGAAATTCCTTGATCTCTTACAGAGGGAAACGATGTACGCCACTGCGC	2436
QY	2334	CGTGAAGTTGACAGTGGTGTGGAAATGAGCACTTCAACATTGACAACTCCAGCTCAAAATTA	2393
Db	2437	AGTGAACCTGTCAGATCCGTGTGGAAACGGCAACTTTGTATTGACAACTCCAGAACTATCCA	2496
QY	2394	AGTTCACTCTACAAAGTGTGGAGCCATGACGTGAAGCTGGGGGTGTGCTCTCAAGCTGA	2453
Db	2497	GGGCACTCTTACAAAGTATCCCGGCCCTGTGGAGAGCTGTGGGCTCTGTGCTTCAAGCGCA	2556
QY	2454	CCCAAGCTTGCATGTGTGGTGTGTCCAGAGGCCAGGCCAAGTGCACCTGTGGCCACACTG	2513
Db	2557	CCCCGCGCTTCAGATGGCGATGTGTGTGGCCGAGCCGGCGCTGCTCTCTGTGGACCAACTG	2616
QY	2514	CCCTGGCCCA---GGAGAGCCAGTGGGTGGAGCTGTGTGTGTGCTCAAAAGTGCACAAA	2570
Db	2617	CGCTTCCGACACACTGTGATGTGGAATGCAAGCCCGCTTCAAGGACAGATGTGTGTGACGA	2676
QY	2571	CCCCCGCATCACAGATATATCCCGTGTACAGSCCCCCGGGAGAGGGGACCAAGTTCAC	2630
Db	2677	CCCCAAGATCTCTAAGCTGTGCCCCCGAGACGGGGCCGAGGAGGCGGACAGCGGCTCAC	2736
QY	2631	TATCCGAGAGGAGAAACCTGGGGCCCGAAATTTGGGACATGCGCTCCCATGTCAAGTTC	2690
Db	2737	TATCACAGGCGAAGACTGGGGCTTGTGCATTTGAAAGACGTGGTCTGGGGCGTGGCGGTGG	2796
QY	2691	TGGCGTGAAGTGCAGCCCTTTAGTGAATGTTACATCCCTTGACAGACAGATCGTGTGTGA	2750
Db	2797	CAAGGTGCTGTGACGCCCTGTGTGAGAGAGGAGTACATCACTGTGCGAGAGATCGTCTGTGA	2856
QY	2751	GATGGGGGAGGCGAAAGCCC---AGCCAGACATGACAGGCTTCGTGAGATCTCGTGGCTGT	2807
Db	2857	GATCGGGGACCGCACTCCGTGTGCTGTGCGCCATGACGCCCTGTGTGAAGTGTGTGTGGGA	2916
QY	2808	GTGTGGGCTGAATTCATGTGAGCCGGTCTCTCAAGCTTATTACTTATGACACTGACTCT	2867
Db	2917	CTGTCAACCAACTCTCCGCGCTGTGTACCCAGAGGCTTCACTTGTGTGACCAACAACCTT	2976
QY	2868	CTCAATCTGAAAGCCACGCGGGGGCCCATGTCCGGAGAGAACCCAAATGACATACACAG	2927
Db	2977	CTACCGTGTGACCCCTCCGTGTGGCTCTGTCAAGGGGACCTGTGAATGTGAATGAGGG	3036
QY	2928	CACCAACTGAATGCGCGAGCAACGCTGTGTGATGTTTGGAAAGACCCCTGTCTTT	2987
Db	3037	AAGCCACTGAACGACAGGACATGATGTGGCTGTGTGCGGTGGCCGCGCTGTCTCTT	3096
QY	2988	CCACAGGGGATCTCCATCTCTACATTTGTGTGCAACCAACATCTCAGATGAGGTGTGAGA	3047
Db	3097	CTCTGTGTCCAGAGGAATCTCCGTGTGAATCCGGTGTCTGACACCCCCGGGACAGGCC	3156
QY	3048	GATGAAGGTGTGCTGTGACAGTGTGACAGAGCC-----AAGATCCACAGAGCTGTGT	3098
Db	3157	TGGCAGCGCTCCCATCTATCATCAACATCAACCGCGCCACGCTCAACAACCTGTGAGTGA	3218
QY	3099	CTTTAGTATGTGGAAGACCCACCACTGTGTGGATGTGAGCCAGATGTGAGCATTTGAG	3158
Db	3217	GTACAACTTACCGAGGACCCCACTCTGTAGATTCGACCCGAGTGTGAGATCAACAG	3276
QY	3159	TGGAAACACACCCATCGCGTATGTGGGGACCACTGTGACCTCATACGAACCCCCAGAT	3218
Db	3277	CGGTGGGACCTCTCTACCGTCAACGGACCAACCTGTGCGCACTGTCTCGTGAACCCCGAAT	3336
QY	3219	CGGTGCAAGCATGTGAGGAGAGGACATCAATATCTGTGAGTCTCTGAAGCTATCTGA	3277
Db	3337	CCGGGCAAGTATGTGAGGCAATTGAG---AGGAGAACTGCTGTGTATCAATGACACAC	3392
QY	3279	GATGACCTGTGAGGCGCGCCCTGTGCTGTGGTCTGTGACACACAGTCAAGACTTGAACGA	3338
Db	3394	CATGTATGTGCGGCGCCCTGTGTGTGCAACCTGTGTGCGAGCCACCAAGACTGTGGGA	3455
QY	3339	GAGGCCGAGGATTTGGCTTCACTCTGACCAAGTCCAGTCCCTCTCATCTCTCAACA	3399

Db	3454	GCGGCCGGAAAGACCTGGGCTCTTCATCAATGACAACGAGCGCCTCCCTGCTGTGCTCAACAC	3513
Qy	3399	GACCAATTCAACCTATCATCCCAACCCGGGTTTAGAGGCTTTGGCTCCCTCAGAACTCT	3458
Db	3514	CACCTCTTCTCTACTACCTTGAACCCCGATACGAGGCACTGACGCCACATGGCCTGCT	3573
Qy	3459	GGAAGCTCAAGCTTGGCAGCGCCCATCATCTTAAAGGGAAGAACTTAATCCCGCTGTGGC	3518
Db	3574	GGAAGCTAAGCCCAAGCTCCCGCACTCATCTCAAGGGCCGGAACCTTGTGCACCT--CG	3630
Qy	3519	TGGGGGCAAGGTGAAGCTGAACATACTGTGCTGTGTGGGAGAAACCGTGACCCGTGAC	3578
Db	3631	ACCGGCAACTCCCGAATTCAACTACAGGGTGTCAATGGGTCCACACCTGTACCTTCAC	3690
Qy	3579	CGTGTCAATGATCCAGCTGCTCTGCGAGTCCCCCAACCTCATGTGCAAGGCAAAATGAT	3638
Db	3691	CGTGTCCGAGACGCAATGCTGTGTCAGAGGGCCCAACCTCATCTGGGAGACAAAGGTCAAC	3750
Qy	3639	GGCCCGTGTGGGTGGCATGAGTAAGTACCCCGGGGATGATGATCAATTGCCCCGACAGCC	3698
Db	3751	GGTGGGTGACAGTGTGCTTCAAGTTCCTGCGAAGGACATGTCAAGTGAATTCGGACAGCT	3810
Qy	3699	GCTCAAGCTGCGCCCATCTGTCAAGATCGCATGTGCGTGGCGCTCTCTCATCAATTTTCA	3758
Db	3811	GCTACAGCTGTCTGCAATTGTGGCATTGGCGAGGCGGGGGGTCTCTGTCTGTGTGAT	3870
Qy	3759	CGTGGCGGTGCTCATTGTACCTATAAAGCAAGTCCCGGGAAGTGACTCACTGTGAAGCG	3818
Db	3871	CGTGGCTGTGCTATCGCTTCAAGCCCAAGTACAGAGATGTGTGACCGGCACTCAAGCG	3930
Qy	3819	GCTCAAGTCAGATGACAACTGTGAGTCCCTGTGGCCCTTGAATGTGCAAGGAAGCTT	3878
Db	3931	GCTGAGCTTCAGATGAGCAACCTGGAGTCCCGGTGGCCTTGAAATGCAAGGAAGCTT	3990
Qy	3879	TGCGAGGTGACAGCGGACATCCATGTAGCTGACAGTGAACCTGGATGAGCCGAGATTC	3938
Db	3991	TGCAGGTGTGACAGCAAGATCCACAGCTGACCAATGACTGGACGGTCCGGCATCC	4050
Qy	3939	GTTCTGTGACTATAGAACTTACACCATTCGGGTGCTTCCAGGAATTGAAGCAACCC	3998
Db	4051	CTTCTTATACACCGGACATATGGCATTCGGGTGCTTTTCTGGGATTCAGAGCAACCC	4110
Qy	3999	TGTCCTCGGAGCCTTGAAGTCCCGGGTACCGGCAAGGCTGTGAGAAAGGCTGAA	4058
Db	4111	TGTGCTCAAGAGATGAGGTACAGGCCAA-----TGTGGAAGATGCTGTAC	4158
Qy	4059	GCTCTTGCCTCAAGTCAACAACAAGGTGTTCTGTGCTTCAATCCGACAGCTGA	4118
Db	4159	ACTGTTCGGGAGCTGCTGACCAAGAAAGCACTTCTGTGACTTCAATCCGACGCTGGA	4218
Qy	4119	GTTCCAGGATGACTTCTCATGCGGAGCCGTGGCAAGTGGCTCACTCATGATGACGT	4178
Db	4219	GGCAACAGCAGCTTCTCTCATGTGGGACCGGGGAATGTGGCTCGCTCATCAGACGGC	4278
Qy	4179	GCTGACAGCAAGCTGAGATAGCCCACTGAATGTCTGAAGAGGTGCTGGCGACCTCAT	4238
Db	4279	CTGTGAGGGCCGAGATGAAATAGCCCAAGCGGTGTCAAGAGAGCTGTCTTCCAGCTCAT	4338
Qy	4239	TGACAAAGACTGTGAGAGCAAGAACCAACCTAAGCTG---CTGCTCAAGAGACTGATC	4299
Db	4339	CGAGAAAGACTGTGAGAGGAAGAAACACCCCAAGTGTACTGGCGCGGCCCAACTGATC	4398
Qy	4296	AGTGGCTGAAGATGCTGACCAATTGTTTACTTCTCTCTAAGAACTTCCTCAAGGA	4355
Db	4399	GGTGGAGAGAAATGCTACTACTGTGTACCTTCTCTTGTATAGTTCTCTAAGGA	4455
Qy	4356	GTTGTGTGGAGAGCCCTTCTCTCTGTCTGTGCTCCATCAAGCAGATGAGAGAGG	4415
Db	4459	GTGGCGTGGGAGCGCGCTGTTATGCTGTGACCGCATCAAGCAGCAGATGAGAGAGG	4518
Qy	4416	CCCATTAAGCCATACGGGGGAGAGCCCGGTAATCCTTGAAGCGAGCAAGACTCATCG	4475
Db	4519	CCCATTTAGCCATACGGGGGAGAGCGCTACTCTCTCTGATGAGCAAGCTCATCG	4575

QY	4476	CCAGCAGATTGACATCAAAAAACCCCGGCTCAGAGTGTGTGACGCCAGACAATATCCAAACAG	4535
Db	4579	GCACGAGATTGACATCAAGACACTGACACCCCTGAACTGTGTGAACCCCTGAAATGAAATG	4638
QY	4536	CCCCGAGTCCCGAATTAAGATCCTCAACTGTGACACCATCACTCTAGGTCAGAGAAAGAT	4595
Db	4539	ACCTGAGAGTGGCGGTGAAAGGGGCTGAGACTGTGACACGSTACACCGAGGCCAAGAGAAAGCT	4698
QY	4596	TCTGAGTGCACCTTCAAGAAATGTGCTTGCTGCCACCGGCCAAGCTCAGATTAATGGA	4655
Db	4639	GCTGAGACGCTGCTTACAAAGGGCGTGGCTCTACTCCCAAGGGCCCAAGGCCGCGGACATGGA	4758
QY	4656	TCTGAGTGTGCGACAAAGAAAGTGGGGCAAGATGATCTTGACAGATTAAGACATTCAC	4715
Db	4759	CCTGAGAGTGGCGCAAGGGCCGACATGAGGCGGATCATCTTCAAGACCAAGACGTTCACAC	4818
QY	4716	CAAGTTTGAATGATTGTGAAGCACTGAACCACTGAGCCCACTACCAAGGTGCAGATGG	4775
Db	4819	CAAGATTGACAAACGATTGGAAGAGGCTGGAACACACTGCTGCTCACTACCAAGGTGACAGCGG	4878
QY	4776	TTCCGTGTGCGAATTAATGTCTCAAGACAGGTGACAGCTTAAACGAGTGAACAACTTCCAC	4835
Db	4879	GTCTCTGGTGGCACACTGGTGGCCAAAGACACTCCGCTTACAACTTCCAACTTCTTCAC	4938
QY	4836	CGTCTCCAGAACCTGACGAATGAATTAATGAATAATGATCCGATACACGGGCAAGCCCGCA	4895
Db	4939	CTTACCAACAG---TCCCTCAGCAGATTAAGAAAGATGTGTGGCAAGGCCAGACCCCGA	4995
QY	4896	CAGCTCCGCTTACGGAACACTATGATCACTCTGACTTGAAGTGTGAATCAAGATGTG	4955
Db	4996	CAGCTTGGCTGGCGCACGGCCACTGATACACCCGACCTTGAAGCGGCACTCAAGCTGTG	5055
QY	4956	GCACCTTAATGAAGAACACAGACACGAGACCAAGAGAGGGGACCGGGGAGACAAAGAT	5015
Db	5056	GCACTCGTGTGAAGAACACAGACCAACCTGGAACAGGTGAGGGTGAACCGCGGACAGAAAT	5115
QY	5016	GATGTCTGAATCTTAAGTCAACCCGACTCTGSCCACTA--AGGGACACTGCAAGATGT	5072
Db	5116	GATCTCGGAAGATCTAATTGACACGGCTACTGSCCAACCAAGACGGGCAACACTGCAAGATGT	5175
QY	5073	TGTGATGACCTCTTTGAGACCATTCTTACGACGGCACACCGTGGCTCTGCCCCCT	5132
Db	5176	TGTGACGACCTGTTTGAGACCATTTTACGACGGCACACCGGGGCTCAGCCCTGCGGCT	5235
QY	5133	GGCCATCAAGTACATGTTTGACTTCTCTGATGAGAGGCTGATTAACATGSCATTTCTGA	5192
Db	5236	GGCCATCAAGTACATGTTTGAATTCCTGATGAGAGGCGGACCAAGCACAGATCTCACGA	5295
QY	5193	CCCGACGTCGSCCATTACCTGGAAAGACAATT--GCTGCCCCGAGAGTTTGGATCAA	5249
Db	5296	TGCTAGCGTGGGCGCACACTGGAAAGACACTGACGCTGCCCCCTCTGGGTCGAA	5355
QY	5250	CATGATCAAGAAACCCGCAATTTGTGTTTGAACATCATTAAGAACAGCATCAAGAGCTCG	5309
Db	5356	CGTGTACAAAGAACCAAGTTTGTGTGCACTTCAAGAACAGCATCAAGAGCTCG	5415
QY	5310	CCTCTCTGTGTGTGCTCAGACCTTCATGGACTTTTGCTCAAGTCAAGACACCGGCTGGG	5369
Db	5416	CTTGTGCGGTGGTGGCCAGACTTTCATGGACTCTGCTCAACCTGTGAGCAAAAGCTGGG	5475
QY	5370	CAAGACCTGCCCCCTCAAAACACTGTGATATCCCAAGAACATCCCCACACTCAAGAAATG	5429
Db	5476	CAGAGACTACCTTCCAAACACTGTGCTTACGCAAGACATCCCACTACAGAGCTG	5535
QY	5430	GGTGGG--GAGGTATTACTCAGACATAGAGAAAGTCCAGCCATCAGCGACCAAGACAT	5486
Db	5536	GGTGGAGAGAGATACTATGACAGACATCGCCAAAGTGCAGCCATCAGCGACCAAGACAT	5595
QY	5487	GAAACGATACCTGGCTGAGACAGATCCCGGATGACATGAATGATGTTCAACACATATGTC	5546
Db	5596	GATGTGATCTGTGCTGAGAGATGCCGCTGCACTTGAACCACTTCAACAGATTAAGGCG	5655

OY		5547	ACTCTGAGAAATCTTCCCAATGAGGCGAAATACAGGA-----GGAGAATCCTTGACC	5600
Db		5656	CTTGACGAAATCTACTCTCAATCATCACCAAGTACAGAGTAGAGGTGCAGATCCTGCCAGC	5715
OY		5601	TCTGACCAAGATGACCAAGTCGTGTGGGAAGCAGAAACTGCGCTTACAACCTAAGAACATGCAT	5660
Db		5716	CCTGAGAGAGATGATGACGAGCGCGCGCCAGCGCTCGAGAGCAAGCTTGAGCAGGTGGT	5775
OY		5661	AACCTCATGAGCTTAAGACAGCTGA	5685
Db		5776	GGACACATGAGCCTTGAGCAGCTGA	5800

RESULT 3

```

US-10-218-779-31
; Sequence 31, Application US/10218779
; Publication No. US20040029222A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit
; APPLICANT: MacDougall, John
; APPLICANT: Millet, Isabelle
; APPLICANT: Ellerman, Karen
; APPLICANT: Stone, David
; APPLICANT: Gerlach, Valerie
; APPLICANT: Grosse, William
; APPLICANT: Alcobrook II, John
; APPLICANT: Lepley, Denise
; APPLICANT: Rieger, Daniel
; APPLICANT: Burgess, Catherine
; APPLICANT: Casman, Stacie
; APPLICANT: Spytek, Kimberly
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Mishra, Vishnu
; APPLICANT: Paturajan, Meera
; APPLICANT: Shenoy, Suresh
; APPLICANT: Rastelli, Luca
; APPLICANT: Tchernenov, Velizar
; APPLICANT: Verner, Corine
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Malyskar, Uriel
; APPLICANT: Guo, Xiaojia
; APPLICANT: Miller, Charles
; APPLICANT: Gangoli, Esha
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-214
; CURRENT APPLICATION NUMBER: US/10/218,779
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: 60/253,834
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/250,-926
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/264,180
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/313,656
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/327,456
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 5895
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-218-779-31

```


Query Match	44.5%;	Score 2533.8;	DB 12;	Length 5895;
Best Local Similarity	68.0%;	Pred. No. 0;		
Matches 3812;	Conservative	0;	Mismatches 1712;	Indels 81;
				Gaps 17

141 CCCCAGGGTTCAATCACTCGTGTGTGATAGAGACACAGACACATTACTTGGGGGC 200

Db 217 CGACTGGGGCTCACCACCCTAGTGTGATGAGCAGACGAGGTGTATGTGGGCGC 276
QY 201 CCGTCAATCGGATTTCAGAGCTCTCCAGCGACCTGMAAGCTTGTGTGAACGATAGACAGG 260
Db 277 AGTGAACCGCATCTATTAAGCTGTGGGAACTGTGACATCTGTGGGGCCCACTGACGG 336
QY 261 GCCGACGAGAGAACCCCAAGTGTACCCACCCCGCATGTGTCCAGACCTGCATAGGCC 320
Db 337 CCGTGTGAAGGACAAAGAAAGTGTACCCGCGCCAGGTGTGACGTCTGCCCCACGG 396
QY 321 CTTGACCAACCAACATGTCAACAGATGTCTCTCATATGACTAAGAGAACAGGT 380
Db 397 CTTGGGCACTACGACAAAGTCAACAGCTGTGTGTGTGTGACTATGCGCTAACGGCT 456
QY 381 GATTGCTGTGGAGCTGTACCAAGGATCTGCAAGCTGTAGAGCTGAGAGCTCTT 440
Db 457 GCTGGCTGTGGAGCGCCCTCCAGGSCATCTGCGATTCTGTGCTGAGATCTCTT 516
QY 441 CAAGCTGGGAGAGCTTTATTAAGAAAGACCTATCTGTCAAGTGTCAAGAGAGCG 500
Db 517 CAACCTGGGTGAGCCACACCAACCGTAAAGGACCTACTGTTCAGGCTGACAGAGGACAG 576
QY 501 CTACGTCTTTGAGTGTATGCTCTCTTACA-----GCAACCTGATGACAAGCTGTTCAT 554
Db 577 CAGCATGGCGGCTGTCTCATTTGCCGGGCCACCGGGCCAGGGCCAGAGCTCTTGT 636
QY 555 TGCCACGAGAGTGTGATGGGAAAGCCGAGTATTTTCCCACTTCCAGCCGAACTGAC 614
Db 637 GGGGACACCCATCGATGGCAAGTCCGAGTACTTCCCACTGTCCAGCCGTGCGTCAAT 696
QY 615 CAABAACCTGAGGCGGAGTGGCATGTTCGCGTACGTCTTCATGATGATGTCTGTGCGTC 674
Db 697 GGGCAAGAGAGAGTGTGCGGACATGTTCGGCTTCGTGTACAGAGATGATTTGTGTATC 756
QY 675 GATGATTAGATCCCTTCGACACACTTACACATCATCTGACTTTGATATCTATGT 734
Db 757 ACAGCTCAAGATCCCTTCGACACGCTGTCCAAATTCGCGCTTTGACATCTACATGT 816
QY 735 CTATGTGTTTGAAGTGTGCACTTGTGTACTTTTGAACCCGTCAACCTGATGTGATC 794
Db 817 GTACAGCTTCGACGACGAGCTGTGTCTACTACCTACGCTGACGTAAACACACAGCT 876
QY 795 TCCACAGGCTCCACACCAAGAGCAGGTGTATATCCAAAGCTGTGAGGCTTTGCA 854
Db 877 GACCTCGCTGATGCGCGCGCGGAGCACTTCTCAGCTCCAAAGTCTGTGCGCTGTGT 936
QY 855 GAGAGACACAGCTTTCATCTCTATGTAGAGTGCCTATGCTGTGAGGCGAGTGGGT 914
Db 937 GAGACACCCCAAAATTCATCTGTAGCTGTGAGTTCCTCATTTGGCTGACAGCGGGGTGT 996
QY 915 GAGAGACAGCTGTGTGACGCTGCTACTGTGTCCAAAGCGGGGCGCTGTGTGACAGAC 974
Db 997 GAGATACCGCTGTGTGACGATGTCTACTGTAGCGCGCGCGCGCTGTGAGGCCACA 1056
QY 975 CTTTGAAGTTCATTCAGATGTATGATGTCTTTCACCGTCTTCTCCAAAGGCGCAAGCG 1034
Db 1057 GCTGGGCTGTGCTGAGCAGAGAGTGTCTGTCTACTGTGTTCGCCAGGCGCAAGAA 1116
QY 1035 GAAATGAAATCCCGATGAGTGTGAGCGGCTGTGACCTTCTACTTTGAAGAGATTAATGA 1094
Db 1117 CCGGTGAAGCACCAAGAGATGACACTGTGTCTGTACGCTGAGGGCAATCAAGGA 1176
QY 1095 CCGCATTAAGAGCGGCTGTGACTGTGTTCACCGGCGAGGCGACGCTGACCTGGCTG 1154
Db 1177 GAAGATTAAAGGCGGCATCCAGTCTGTACCGTGTGAGGGAAGCTTCTCCGCGTG 1236
QY 1155 GCTCAAGGTGAAGGACATCCCTGTGACAGTGCCTTAAACATTAAGATTAATCTG 1214
Db 1237 GCTGTCAACAAGAGCTGGGCTGTATACCTGCGCTGTGACATGATGACATCTTG 1296
QY 1215 TGGCTGTGACATGAATGTCTCCCTGGAGGTGTCCGACATGTGTGTGAATTTCCGCTT 1274
Db 1297 CGGGCAGGACTTCAACGAGCCCTGGGGGCGACAGTCAACATTAAGGGAAGCGCCCTGT 1356

QY 1275 CACGAGAGACAGGACCGCATGACGCTGTATCATGCATATGTGTACAAAGACACTCTCT 1334
Db 1357 CGTGGAACAAGATGATGGCTTGACCGCGCTGTGCTGTATGATATCGGGGCGCACTGT 1416
QY 1335 GCGCTTGTGGGACCAAAAGTGTGCAAGCTGAAGAAAGATCCGGTGTGATGAGCCCA----- 1390
Db 1417 GGTATTCGCGGACCGGAAAGTGTGCGCATCCGCAATCTGTGTGACCTCTCCAAACC 1476
QY 1391 -----GGGGCAACGCTCCAGATATGAGACGATGTGCAAGGTGTGACCCCGGCACTGT 1445
Db 1477 CGGTGGCGGCTGTGCTGTGCGCTACGAGACGCTGTGGCCAGAGAGGGGAGCCCACTCT 1536
QY 1446 CCGGATATGCGCTTCTCCAGAGACCAAGACCACTTCAATGATGATGACAGAGGACGT 1505
Db 1537 GCGAGACCTGTCTCAGCGCCCAACCAAGTACTCTTACGCAATGACCAAGAGACAGT 1596
QY 1506 CACCAAGTCCCTGTGAGTCTGTGTGATGATACAGAGCTGTGCGGACAGTCTTGGCTC 1565
Db 1597 GACCGGCTGTCTGTGAGAGCTGTGTGTGACGTACCTCTGTGAGCTGTGTGCGGGTCT 1656
QY 1566 AGGGAACCCCACTGT 1625
Db 1657 ACGGGACCCCACTGT 1716
QY 1656 TGAGCGGTCCAAGAGCCCGCAGATTGTGCTCGGAGATGAGACAGTGTGTCCGGCTGAC 1685
Db 1717 TGAGGAGACACAGACCCCAAGCCCTTGTGTGCGGACCTGTGTGAGTGTGTGACGTGAC 1776
QY 1686 GATTCATCCACAAATATCTCTGTCTCTAGTACACAGTGTG-----TGTGTGTCTGTGAGAC 1742
Db 1777 TGTGAGACCCCGCAATGT 1836
QY 1743 GTACAAATGTCCCGGAGCTGTGAGCTGTGAGCTGTGACCTTTGAGAGACCTGTGAGAT 1802
Db 1837 CTGHAAGCTGTGTGACCTCTGTAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1896
QY 1803 GATGTGGT 1862
Db 1897 TGAGAGCTCTGTGAGATGTGCTCGGATCTACGTGTGTGTGTGTGTGTGTGTGTGTGT 1956
QY 1863 CCGATCA-----TCAAGAGAAATGGGGAACCAACATGTGTGTGTGTGTGTGTGTGTGT 1916
Db 1957 GCGCATACGCGGGGCGAGGTGTGAGAGACCAAGGAGTGTGAATCTTAAAGTGT 2016
QY 1917 AAAGGAGACCGGCATGACCTTGTGCGACGACGCTTGTGTGTGTGTGTGTGTGTGTGTGT 1976
Db 2017 CAAGGAGACAGGGAAGATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2076
QY 1977 CAATTG---TGCCTGTCTGT 2033
Db 2077 CCAATGTGAGCTGT 2136
QY 2034 GCATGTGTGACCCCAATGACCCCAAGACCTGTCTTCTTCCAGAGAGCGAGTGAAGCTGCG 2093
Db 2137 CCAAGTGTGACACACCAAGT 2196
QY 2094 CAGAGATGTGCCCCAGCTGT 2153
Db 2197 TGAGGACTGTCCCAAGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2256
QY 2154 GCTATACAGCTGAAGGCGCAAGAACCTCCCAAGCCCAAGTGTGTGTGTGTGTGTGTGTGT 2213
Db 2257 ACCATACCTGT 2316
QY 2214 ATGATCTCAACATTCAGGGGTGAGAGAGAGATGCGCGCTGTGTGTGTGTGTGTGTGTGT 2273
Db 2317 GTGCTCTTCCACATTCGCGGAGCGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2376
QY 2274 CAGGATAGTGTGCAAGAACACTTATTTCTATGAGAGATGAGATCAACAACTGTGCG 2333
Db 2377 CAGCTGAGTGTGCAAGATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2436

QY 2334 CGTGGAGTTGACGCTGCTGTGGAAATGGGCACTTCAATTGACCAACCCAGCTCAGAAATA 2393
Db 2437 AGTGAACCTGTCACTGTGTGGAAACGGCACTTTGTCTATTTGACACCCACAGAACTATCA 2496
QY 2394 AGTTCACTTCTACAACTGTGGAAACCTGCTGAGAGCTGCGGCTGTGCTTCAAGCTGA 2453
Db 2497 GGGCAACCTCTTCAAGAGCCCGGCTGCGCGAGAGCTGCGGCTGTGCTTCAAGCTGA 2556
QY 2454 CCCAGACTTGGCAATGTGCTGTGGAGCCAGAGGCCAGAGCCAGTGAACCTGCGGCGAGACTG 2513
Db 2557 CCGCGCTTGAAGTGGAGTGTGCTGTGCGGAGCGCGCTGTCTCTGCGACACCACTG 2616
QY 2514 CCGTGGCCCA--GGAGAGCCAGTGTGCTGAGCTGTGTGCTGTGCGMAAAGCAAGTGCACAA 2570
Db 2617 CGGTGCGGACACACTGTCATCTGTGAATGACAGCGCGCTCAAGGAGAGTGTGCTGACCGA 2676
QY 2571 CCCCCGATCAAGAAATAATCCCGGTGACAGGCCCGCGGAGAGGGGGGACCAAGTGCAC 2630
Db 2677 CCCCAGATCTTAAGCTGTCCCGGAGAGCGGCGGAGCGGAGCGGCTCAC 2736
QY 2631 TATCCGAGGGAGAACTTGGGCTGTGAATTTCGCGACATGCGCTTCCATGTCAAGGTTGC 2690
Db 2737 TATACAGGGAGAACTTGGGCTGTGAGATTGMAAGGTGCGCTGTGCGGCTGTGCGGCTGAG 2796
QY 2691 TGGCTGTGAGTGCAGCCCTTTAATGATGTTTACATCCCTGCAAGAACAGATGTGTGTA 2750
Db 2797 CAAGTGTGTGAGCCCTGTGAGAGCGAGTACATCAGTGCAGAGCAAGTGTGTGTGTA 2856
QY 2751 GATGGGGAGAGCCCAAGCCC--AGCCAGCATGCAAGGCTTGTGAGATCTGCTGTGCTGT 2807
Db 2857 GATGGGGAGAGCCAGCTCCGCTGCTGTGCGCTGACAGCGCTGTGAGATGTGTGTGCGGA 2916
QY 2808 GTGTGCGCTGAATTCATGTCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 2867
Db 2917 CTGTCTACACACTACCGCGCCCTGTGACCCAGCGCTTCACTTGTGTGACACCAACTT 2976
QY 2868 CTCAATCTGAAGCCAGCGGGGGGCCATGTCGGAGGGAGCCCAAGTGCATCTACAG 2927
Db 2977 CTACCTGTGAGCCCTTCCGTTGCGGCTGTGTGTGAGGGGACCTGTGATGTGACATCAAG 3036
QY 2928 CACCAACTGTAATGCCGAAGCAAGTGTGTGATGTTTGAAGAGCGCTGTCTCTT 2987
Db 3037 AAGCACTGTGAACCAAGGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3096
QY 2988 CCAAGCGGATCTCATCTTCAATTTGTGTGCAACACATCTCTCAATGAGTGTGTGA 3047
Db 3097 CTCTGTGTGAGAGAACTCCGTTGATTCGGGTCTTGAACCCCCCGGGAGAGCCC 3156
QY 3048 GATGAAGGTGTGCTGTGAGGTGACAGGGCC-----AGATTCACCAAGCACTGTGT 3098
Db 3157 TGGAGAGGCTTCCATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 3216
QY 3099 CTTTCAATGTGTGAAGACCCACATCTGTGCGATTTGAGCGAGATGTGAGCTTTGTGAG 3158
Db 3217 GTACAACTTACACCGAGAGCCCACTGTGAGATGAGCCCGAGTGGAGCATCAAG 3276
QY 3159 TGAAGAACCAACCTGCGGTATGGGGGACCACTGTGAGCTTCAATGAGAACCCCGAGAT 3218
Db 3277 CCGTGGAGACCTTCTTGAAGGTGACAGGCAACCACTGTGCGGTATGAGAACCCCGAAT 3336
QY 3219 CCGTGGAGCATGAGAGGAGAGAGCAATCAATATCTGTGAGTTTGAAGCTTACTGA 3278
Db 3337 CCGGCGCAAGTATGAGAGCATTTAG--AGGAGAACTGTGCTGTGTATGATGACACAC 3393
QY 3279 GATGACCTGTGAGGCGCGCTGTGCTGTGCTGTGCTGTGACCAAGTCAAGCTGACGA 3338
Db 3394 CATGTATGCGCGCGCTGT 3453
QY 3339 GAGGCGCGAGAGTTTGTGCTTCACTGTGACCAAGTGTGCTGTGCTGTGCTGTGCTGTG 3398
Db 3454 GCGGCGGATGAGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3513
QY 3399 GACCACTTCACTACTATCCCAACCGGCTGTGTGAGGCTTTGTGTGTGTGTGTGTGTGT 3458

Db 3514 CACCTCTTCTCTTACTATACCTGTGACCCGATGAGAGCCACTGAGCCCCACTGGCTGTCT 3573
QY 3459 GGAAGTCAAGCTGTGAGAGCCGCTTCAATCTTAAAGGGAGAAACCTGTATCCGCTGTGAG 3518
Db 3574 GGAAGTCAAGCTGTGAGAGCCGCTTCAATCTTAAAGGGAGAAACCTGTGTGCTGTGCT 3630
QY 3519 TGGGGGCAAGCTGAAAGCTGAACTACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3578
Db 3631 ACCGGGAACTCCGAGCTCAATCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3690
QY 3579 CGTGTCAAGATGTCAAGT 3638
Db 3691 CGTGTGAGAGAGCAAGT 3750
QY 3639 GAGCGGTGTGCTGATGAGT 3698
Db 3751 GGTGTGTGAGT 3810
QY 3699 GCTGAGCTGTGCGCATGT 3758
Db 3811 GCTGAGCTGTGCTGT 3870
QY 3759 GGT 3818
Db 3871 GGT 3930
QY 3819 GCTGTGAGT 3878
Db 3931 GCTGTGAGT 3990
QY 3879 TGGCGAGTGTGAGAGCGGATCTTCAATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3938
Db 3991 TGCAGAGTGTGAGAGCGGATCTTCAATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4050
QY 3939 GTTCTGTGACTATGAACTTACACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3998
Db 4051 CTTCCTGTGACTATGAGCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4110
QY 3999 TGTCTTCCGGAGCTTGT 4058
Db 4111 TGTGTCTAAGAGT 4158
QY 4059 GCTTGTGCGGAGCTTCAATCAACCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4118
Db 4159 ACTGTGTGGGAGAGCTGT 4218
QY 4119 GTTCCAGGCTGT 4178
Db 4219 GGCACAGGAGCTTGT 4278
QY 4179 GCTGTGAGAGAGT 4238
Db 4279 CTTGTGAGGCGAGATGT 4338
QY 4239 TGAAGAACTGT 4295
Db 4339 CGAGAGAACTGT 4398
QY 4296 AGTGTGTGAGAGT 4355
Db 4399 GGTGTGTGAGAGT 4458
QY 4416 GGTGTGTGAGAGT 4475
Db 4519 CCGCAATGAGCATCAAGGAGGAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4578
QY 4476 CAGAGAGTGTGACTCAAAACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4535

Db 4579 GCAGCAGATTGACTACAAGACACTGACCCTGAACTGTGTGAACCTTGAAATGAAATGC 4638
Qy 4536 CCCCAGGTCCCAAGTAAGATTCCTCACTGTGACACCACTCACTCAAGTCAAGGAAAGAT 4595
Db 4639 ACCTAGAGTGCCTGTAAGAGGCTGACTGTGACACCGGTCACTCCAGGCGCAAGAGAACT 4658
Qy 4596 TCTGATGCGCATCTTCAAGATATGTCTCTCCACCGGCGCAAGCTGCAATATGA 4655
Db 4699 GTTGACCGCTGCTCAAGAGGCTGTCTTCTCCACCGGCGCAAGCTGCAATATGA 4758
Qy 4656 TCTGATGTCGCAACAAGAGTGGGCAAGATGATCTTGACAGATGTAAGATCAACAC 4715
Db 4759 CTTGAGTGGCGCCAGGCGCGCATGCGGCACTCACTTGACAGAGAGCTCACAC 4818
Qy 4716 CAAGATTGAGATTTGAGAGCGACTGAAACAACCTGGCCCACTACAGTGGCCAGATG 4775
Db 4819 CAAGATTGACACGATTTGAGAGAGCTGAAACAACCTGGCTCACTACAGTGGAGACAG 4878
Qy 4776 TTTCCGTGTGGCATTTAGTGTCCAGAGGTGACAGCTTTAAAGCAATGTAACACTCAC 4835
Db 4879 GTCTTCGTGTGGCACTGTGTCCAGAGCAAGTCCGCTTCAACATCTCCACTCTCAC 4938
Qy 4836 CCGTCCAGAGCCTGACAGATTAATATGAAACAATGATCCGCTTCAACAGGAGCCCGA 4895
Db 4939 CTTTACCAAG--TTCCCTCAGCAGATTCAGAGCATGTCTGGCCACAGCCACAGCCCGA 4995
Qy 4896 CAGCCTCGCTCAAGGACACTATGATCACTCTTGAAGTGTGAGTCAAGATGTG 4955
Db 4996 CAGCCTCGCTCGCGACCCCATGATCAAGCCCACTGAGAGCAGGACCAACCTGTG 5055
Qy 4956 GCACTTATGAGAAACAAGAGAGGAGACCAAGAGAGGAGGAGCCGAGGAGAGCAAT 5015
Db 5056 GCACTTATGAGAAACAAGAGAGGAGACCAAGAGAGGAGGAGGAGGAGGAGGAGGAG 5115
Qy 5016 GGTGTGTGAATTTACCTGACCCGACTCTGTGCACTA--AGGCACTGTGCAAGTT 5072
Db 5116 GGTGTGTGAATTTACCTGACCCGACTCTGTGCACTAAGGCACTGTGCAAGTT 5175
Qy 5073 TGTGATGACCTTTTGAAGCATTTTGAAGAGGACACGAGTGTGAGTGTGAGTGTGAGTGTG 5132
Db 5176 TGTGAGAGCTTTTGAAGCATTTTGAAGAGGACACGAGTGTGAGTGTGAGTGTGAGTGTG 5235
Qy 5133 GGCATCAAGTACATTTTGAAGTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 5192
Db 5236 GGCATCAAGTACATTTTGAAGTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 5295
Qy 5193 CCGGCACTGCGCATACCTGTAAGAGCAATT--GCTGCTGCTGAGTGTGAGTGTGAGTGTG 5249
Db 5296 TGTGAGTGTGCGCACTGTAAGAGCACTGTAAGAGTGTGAGTGTGAGTGTGAGTGTG 5355
Qy 5250 CATGATCAAGACCCGAGTGTGATGTAAGTGTGATGTAAGTGTGATGTAAGTGTGATGTAAG 5309
Db 5356 CCGTATCAAGACCCGAGTGTGATGTAAGTGTGATGTAAGTGTGATGTAAGTGTGATGTAAG 5415
Qy 5310 CTTCTGTGTGTGTGAGTGTGATGTAAGTGTGATGTAAGTGTGATGTAAGTGTGATGTAAG 5369
Db 5416 CTTCTGTGTGTGTGAGTGTGATGTAAGTGTGATGTAAGTGTGATGTAAGTGTGATGTAAG 5475
Qy 5370 CAAGGATGCGCCGCAAGAGTGTGATGTAAGTGTGATGTAAGTGTGATGTAAGTGTGATGTAAG 5429
Db 5476 CAAGGATGCGCCGCAAGAGTGTGATGTAAGTGTGATGTAAGTGTGATGTAAGTGTGATGTAAG 5535
Qy 5430 GGTGGA--GAGTATTACTCAAGATGTAAGTGTGATGTAAGTGTGATGTAAGTGTGATGTAAG 5486
Db 5536 GGTGAGAGAGGATGTAAGTGTGATGTAAGTGTGATGTAAGTGTGATGTAAGTGTGATGTAAG 5595
Qy 5487 GAGAGCATCTGCTGAGAGTGTGATGTAAGTGTGATGTAAGTGTGATGTAAGTGTGATGTAAG 5546
Db 5596 GAGTGTGTGTGTGTGAGTGTGATGTAAGTGTGATGTAAGTGTGATGTAAGTGTGATGTAAG 5655
Qy 5547 ACTCTCAGAGTCTTCTCTTGTGAGCAATAAGAG--GAGATCTTGTGAGTGTGAGTGTGAGTGTG 5600
Db 5656 CTTTCAAGAGTCTTCTCTTGTGAGCAATAAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 5715

Qy 5601 TCTGACACAGATGACCAAGTGTGAGAGAGCAAACTGGCTTAACTAAGAACTAGTCTAT 5660
Db 5716 CTTGAGAGAGATGACAGAGCGCGCGCAGCGCTGCGAGACCAAGTGTGAGAGAGTGTGT 5775
Qy 5661 AACCTTATGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 5685
Db 5776 GAGACAGATGGCTTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 5800

RESULT 4
US-10-108-260A-802
Sequence 802, Application US/10108260A
Publication No. US20040005560A1
GENERAL INFORMATION:
APPLICANT: HELIX RESEARCH INSTITUTE
TITLE OF INVENTION: No. US20040005560A1 full length cDNA
FILE REFERENCE: H1-A0106
CURRENT APPLICATION NUMBER: US/10/108,260A
NUMBER OF SEQ ID NOS: 5458
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 802
LENGTH: 3666
TYPE: DNA
ORGANISM: Homo sapiens
US-10-108-260A-802

Query Match 20.3%; Score 1157.8; DB 15; Length 3666;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1159; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4527 TGCCAAAGCCCGAGGTCCAGTAAAGATCCTCAACTGTGACCATCACTAGGTCTCA 4586
Db 1 TGCCAAAGCCCGAGGTCCAGTAAAGATCCTCAACTGTGACCATCACTAGGTCTCA 60
Qy 4587 GAGAGATTTCTGATGATGATCTTCAAGATTTGCTTGTCTCCACCGGCGCAAGTGTG 4646
Db 61 GAGAGATTTCTGATGATGATGATCTTCAAGATTTGCTTGTCTCCACCGGCGCAAGTGTG 120
Qy 4647 AGATATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4706
Db 121 AGATATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
Qy 4707 CATCACCACCAAGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4766
Db 181 CATCACCACCAAGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
Qy 4767 GCCAGATGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4826
Db 241 GCCAGATGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
Qy 4827 CAATCCACCGTCTCCAGAGCTTCAAGAGTAAATATGAAACATGATCCGTTACACGAG 4886
Db 301 CAATCCACCGTCTCCAGAGCTTCAAGAGTAAATATGAAACATGATCCGTTACACGAG 360
Qy 4887 CAGCCCGGACAGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4946
Db 361 CAGCCCGGACAGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
Qy 4947 CAAGATGAGCACTTGTGAAAGCAACAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAG 5006
Db 421 CAAGATGAGCACTTGTGAAAGCAACAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAG 480
Qy 5007 GAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5066
Db 481 GAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
Qy 5067 GAGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5126
Db 541 GAGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
Qy 5127 GCCCTGCGCATCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5186

Db	601	GGCCCTGGCCATCAAGTACATGTTTGAATCTTCTGATATGACAGGCTGATTAACATGGCAT	660
QY	5187	TCATGACCCCGACCTGCCCATCTCTGGAAGACAAATTGCTGCCCTTGAGGTTTGGCT	5248
Db	661	TCATGACCCCGACCGCCGCATCTGGAAGCAATTGCCGCCCGAGGTTTGGCT	720
QY	5247	CAACATGATCAAGAACCCGACAGTTGTGTGATCATCATTAAGAACAGCATCAACAGCGC	5308
Db	721	CAACATGATCAAGAACCCGACAGTTGTGTGATCATCATTAAGAACAGCATCAACAGCGC	780
QY	5307	CTGCCCTCTCTGTGTGTGCTCAGACCTTCATGACCTCTTGCTTCAGAGTCAAGACCCGCT	5368
Db	781	CTGCCCTCTCTGTGTGTGCTCAGACCTTCATGACCTCTTGCTTCAGAGTCAAGACCCGCT	840
QY	5367	GGGCAAGACCTGCGCTCCCAACAAAGCTGTGATGCCAAGACATCCCGAGTCAAGAA	5428
Db	841	GGGCAAGACCTGCGCTCCCAACAAAGCTGTGATGCCAAGACATCCCGAGTCAAGAA	900
QY	5427	TTGGGTGAGAGGTATTACTCAGACATAGGGAAGTGCACGCATCAGGCATCAAGACAT	5488
Db	901	TTGGGTGAGAGGTATTACTCAGACATAGGGAAGTGCACGCATCAGGCATCAAGACAT	960
QY	5487	GAACGCATACCTGGCTGACAGAGTCCCGATGACATGATGATGAGTTCAACACCATGAGTC	5548
Db	961	GAACGCATACCTGGCTGACAGAGTCCCGATGACATGATGATGAGTTCAACACCATGAGTC	1020
QY	5547	ACTCTCAGAGATCTTCTCTTANTGTGGGCAATACAGCGAGAGATCTTTGACCTCTGGA	5608
Db	1021	ACTCTCAGAGATCTTCTCTTANTGTGGGCAATACAGCGAGAGATCTTTGACCTCTGGA	1080
QY	5607	CCACATGACCAAGTGTGGGAACCAAACTGGCTTCAAACTAGAACAAAGTCATTAACCT	5668
Db	1081	CCACATGACCAAGTGTGGGAACCAAACTGGCTTCAAACTAGAACAAAGTCATTAACCT	1140
QY	5667	CATGAGCTTAGCAGCTGAAA	5687
Db	1141	CATGAGCTTAGCAGCTGAGA	1161
RESULT 5			
US-10-245-103-91			
Sequence 91, Application US/10245103			
Publication NO. US20030068778A1			
GENERAL INFORMATION:			
APPLICANT: Baker, Kevin			
APPLICANT: Eaton, Dan			
APPLICANT: Filvaroff, Ellen			
APPLICANT: Goddard, Audrey			
APPLICANT: Grimaldi, J. Christopher			
APPLICANT: Guiney, Austin			
APPLICANT: Smith, Victoria			
APPLICANT: Stephan, Jean-Phillippe			
APPLICANT: Watanabe, Colin			
APPLICANT: Wood, William			
APPLICANT: Zhang, Zemin			
APPLICANT: Fong, Sherman			
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC			
FILE REFERENCE: P3630R1C112			
CURRENT APPLICATION NUMBER: US/10/245, 103			
CURRENT FILING DATE: 2002-09-17			
PRIOR APPLICATION NUMBER: 10/197942			
PRIOR FILING DATE: 2002-07-18			
PRIOR APPLICATION NUMBER: 60/059114			
PRIOR FILING DATE: 1997-09-17			
PRIOR APPLICATION NUMBER: 60/063046			
PRIOR FILING DATE: 1997-10-24			
PRIOR APPLICATION NUMBER: 60/065027			
PRIOR FILING DATE: 1997-11-10			
PRIOR APPLICATION NUMBER: 60/079689			
PRIOR FILING DATE: 1998-03-27			
PRIOR APPLICATION NUMBER: 60/086478			

```

      PRIOR FILING DATE: 1998-05-22
      PRIOR APPLICATION NUMBER: 60/087607
      PRIOR FILING DATE: 1998-06-02
      PRIOR APPLICATION NUMBER: 60/089801
      PRIOR FILING DATE: 1998-06-18
      PRIOR APPLICATION NUMBER: 60/090557
      PRIOR FILING DATE: 1998-06-24
      PRIOR APPLICATION NUMBER: 60/090689
      PRIOR FILING DATE: 1998-06-25
      Remaining Prior Application data removed - See File Wrapper or PALM.
      NUMBER OF SEQ ID NOS: 116
      SEQ ID NO: 91
      LENGTH: 2597
      TYPE: DNA
      ORGANISM: Homo Sapien
      US-10-245-103-91

Query Match      13.9%; Score 793.8; DB: 14; Length 2597;
Best Local Similarity 99.7%; Pred. No. 1.6e-215;
Matches 795; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

      QY 579 CGAGTATTTCCCAACCATCTCCAGCCGGAAACTGACCAAGAACTCTGAGCGGATGGCAT 638
      DB 1 CGAGTATTTCCCAACCATCTCCAGCCGGAAACTGACCAAGAACTCTGAGCGGATGGCAT 60
      QY 639 GTTGGCGGAGCTCTCCCAATGATAGTTGCTGGGCTCGATGATTAAGATCCCTTGAGAC 698
      DB 61 GTTGGCGGAGCTCTCCCAATGATAGTTGCTGGGCTCGATGATTAAGATCCCTTGAGAC 120
      QY 699 CTTCACCATCATCCCTGACTTTGATATATCTATATGCTATAGTCTTTAGCAATGGCACTT 758
      DB 121 CTTCACCATCATCCCTGACTTTGATATCTATATGCTATAGTCTTTAGCAATGGCACTT 180
      QY 759 TGTCTACTTTTGAACCTCCCAACTGAGATGCTGTCTCCACGAGCTCCACCAAGAA 818
      DB 181 TGTCTACTTTTGAACCTCCCAACTGAGATGCTGTCTCCACGAGCTCCACCAAGAA 240
      QY 819 GCAAGGTATATCATCCAAAGCTCGTAGGCTTTGCAAGAGGACACAGCTTCAACTCTCTA 878
      DB 241 GCAAGGTATATCATCCAAAGCTCGTAGGCTTTGCAAGAGGACACAGCTTCAACTCTCTA 300
      QY 879 TGTAGAGGTGCCATTGGCTGTGAGCGCAAGTGGGCTGAGTACCGCTGTGCAAGCTGC 938
      DB 301 TGTAGAGGTGCCATTGGCTGTGAGCGCAAGTGGGCTGAGTACCGCTGTGCAAGCTGC 360
      QY 939 CTACTCTGTCCAAAGCGGGGCGCGTCTTGGAGAGACCTTGGAGTCCATCCAGATGATGA 998
      DB 361 CTACTCTGTCCAAAGCGGGGCGCGTCTTGGAGAGACCTTGGAGTCCATCCAGATGATGA 420
      QY 999 CCTGCTCTTCAACGCTCTTCTCCAAAGGGCCAGAAACGGAATAAGAAATCCCTGATGAGTC 1058
      DB 421 CCTGCTCTTCAACGCTCTTCTCCAAAGGGCCAGAAACGGAATAAGAAATCCCTGATGAGTC 480
      QY 1059 GGCCCTGTGACATCTTCACTTGAAGCAGATTAAGACCGCATTAAGAGCGGGCTGCAAGTC 1118
      DB 481 GGCCCTGTGACATCTTCACTTGAAGCAGATTAAGACCGCATTAAGAGCGGGCTGCAAGTC 540
      QY 1119 TTGTATACCGGAGGACAGGACGCTGACCTTGCCCTGAGCTCAAGGTGAAGACATCCCTG 1178
      DB 541 TTGTATACCGGAGGACAGGACGCTGACCTTGCCCTGAGCTCAAGGTGAAGACATCCCTG 600
      QY 1179 CAGAGAGCGGCTTAACAATTGACGATTAATCTTGCTGGGCTTGACATGAATGCTCCCT 1238
      DB 601 CAGAGAGCGGCTTAACAATTGACGATTAATCTTGCTGGGCTTGACATGAATGCTCCCT 660
      QY 1239 GGAAGTGTCCGACATGTGTGCTGGAATTCCTCGTCTTCAACGAGACAGAGACCGCATGAC 1298
      DB 661 GGAAGTGTCCGACATGTGTGCTGGAATTCCTCGTCTTCAACGAGACAGAGACCGCATGAC 720
      QY 1299 GTCTGATCATGGCATATGTCTAACAAAGACCACTCTCTGGCCTTTGTGGGCAACCAAGTGG 1358
      DB 721 GTCTGATCATGGCATATGTCTAACAAAGACCACTCTCTGGCCTTTGTGGGCAACCAAGTGG 780

```

Mon Feb 23 09:34:02 2004

us-09-964-956-12.rnpb

Page 12

QY	1359	CAAGCTGAAGAGATCC	1375
Db	781	CAAGCTGAAGAGGTGC	797

```

RESULT 6
US-10-245-107-91
: Sequence 91, Application US/10245107
: Publication No. US2003006872A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Eaton, Dan
APPLICANT: Filvaroff, Ellen
APPLICANT: Goddard, Audrey
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin
APPLICANT: Smith, Victoria
APPLICANT: Stephan, Jean-Philippe
APPLICANT: Wacande, Colin
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
APPLICANT: Fong, Sherman
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACTS ENCODING THE SAME
FILE REFERENCE: P3630R1C71
CURRENT APPLICATION NUMBER: US/10/245,107
CURRENT FILING DATE: 2002-09-16
PRIOR APPLICATION NUMBER: 10/197942
PRIOR FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/059114
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/063046
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/065027
PRIOR FILING DATE: 1997-11-10
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/086478
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090689
PRIOR FILING DATE: 1998-06-25
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 116
SEQ ID NO 91
LENGTH: 2597
TYPE: DNA
ORGANISM: Homo Sapien
US-10-245-107-91

```

	Query Match	Similarity	13.9%	Score 793.8	DB 14	Length 2597
	Best Local	Similarity 99.7%	Pred. No. 1.6e-215	Mismatches 2	Indels 0	Gaps 0
	Matches 755	Conservative	0			
Qy	579	CGAGTATTTTCCACCATCTCCAGCCGGAATCTGACCAAGAACTCTGAGCGGATGCGAT	638			
Db	1	CGAGTATTTTCCACCATCTCCAGCCGGAATCTGACCAAGAACTCTGAGCGGATGCGAT	60			
Qy	639	GTTGCGGTAAGCTTTCCATGATAGTTCGCGCTCGATGATTAAGATCCCTTCGGACAC	698			
Db	61	GTTGCGGTAAGCTTTCCATGATAGTTCGCGCTCGATGATTAAGATCCCTTCGGACAC	120			
Qy	699	CTTCACCAATCATCCCTGACGTTTGATATCTAATATGCTATGTTTAAAGATGGCAACT	758			
Db	121	CTTCACCAATCATCCCTGACGTTTGATATCTAATATGCTATGTTTAAAGATGGCAACT	180			
Qy	759	TGCTACTTTTGTACCTCCAACTAGATGTGTCTCCACCAAGGCTCCACCAAGGA	818			

D	b	181	TGTCCTACTTTTGAACCTCCAACTGAGATGATGTCTCCACAGAGCTCCACACCAAGGA	240
Q	y	819	GCAGGTGTATACATCCAAAGCTCGTAGGCTTTTGCAGAGAGACACAGACCTTCAACTCTTA	878
D	b	241	GCAGGTGTATACATCCAAAGCTCGTAGGCTTTTGCAGAGAGACACAGACCTTCAACTCTTA	300
Q	y	879	TGTAGAGGTCCCATTTGGCTGTGAGCGCAGTGGGGGTGAGATACCGCTGTGACAGCTGC	938
D	b	301	TGTAGAGGTCCCATTTGGCTGTGAGCGCAGTGGGGGTGAGATACCGCTGTGACAGCTGC	360
Q	y	939	CTACCTGTCCAAAGCGGGGCGCGTCTTGGCAGAGACCTTGGAGTCCATCCAGATGTA	998
D	b	361	CTACCTGTCCAAAGCGGGGCGCGTCTTGGCAGAGACCTTGGAGTCCATCCAGATGTA	420
Q	y	999	CTGTCTCTTCAACGCTCTTCTTCCAAAGGGCCAGAAACGGAAATGAATCCCTGATGATGTC	1058
D	b	421	CTGTCTCTTCAACGCTCTTCTTCCAAAGGGCCAGAAACGGAAATGAATCCCTGATGATGTC	480
Q	y	1059	GGCCCTGTGATCTTCACTTTGGAAGCAGATPAAATGACCGCATTTAAGAGCGGCTGCAGTC	1118
D	b	481	GGCCCTGTGATCTTCACTTTGGAAGCAGATPAAATGACCGCATTTAAGAGCGGCTGCAGTC	540
Q	y	1119	TTGTAACTGGGGGCGAGGGACAGCTGACCTTGACCTGGCTCAAGGTGAAGACATCCCTG	1178
D	b	541	TTGTAACTGGGGGCGAGGGACAGCTGACCTTGACCTGGCTCAAGGTGAAGACATCCCTG	600
Q	y	1179	CAGCAGTGGCTCTTAAACAATTGACGATTACTCTTGAGGCTGGACATGATGCTCCCT	1238
D	b	601	CAGCAGTGGCTCTTAAACAATTGACGATTACTCTTGAGGCTGGACATGATGCTCCCT	660
Q	y	1239	GGAGGTGTCCACATGGTGTGAGTGTGATTTCCGCTTCAACGAGGACAGGACCGCATGAC	1298
D	b	661	GGAGGTGTCCACATGGTGTGAGTGTGATTTCCGCTTCAACGAGGACAGGACCGCATGAC	720
Q	y	1299	GTCGTGTCATGCGATATGCTTACAGAAACCACTCTCTGGCTTTGTGGGACCAAAAGTG	1358
D	b	721	GTCGTGTCATGCGATATGCTTACAGAAACCACTCTCTGGCTTTGTGGGACCAAAAGTG	780
Q	y	1359	CAAGCTAAGAGAATCC 1375	
D	b	781	CAAGCTAAGAGAATGTC 797	

RESULT 7
US-10-245-143-91
Sequence 91, Application US/10245143
Publication No. US20030068780A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Bacon, Dan
APPLICANT: Filvaroff, Ellen
APPLICANT: Goddard, Audrey
APPLICANT: Grunwaldi, J. Christopher
APPLICANT: Gurney, Austin
APPLICANT: Smith, Victoria
APPLICANT: Stephan, Jean-Philippe
APPLICANT: Watande, Colin
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Fong, Sherman
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3630R1C90
CURRENT APPLICATION NUMBER: US/10/245,143
CURRENT FILING DATE: 2002-09-16
PRIOR APPLICATION NUMBER: 10/197942
PRIOR FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/059114
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/065046
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/065027
PRIOR FILING DATE: 1997-11-10

QY 759 TGTCTACTTTTGGACCTTCCAACTGAGATGCTGTCTCCACCGAGCTCCACCAAGGA 818
DB 181 TGTCTACTTTTGGACCTTCCAACTGAGATGCTGTCTCCACCGAGCTCCACCAAGGA 240
QY 819 GGAGGTGTATACATCCAAAGCTGTGAGGCTTTGCAAGAGACACAGCTTCAACTCTTA 878
DB 241 GCAAGGTGTATACATCCAAAGCTGTGAGGCTTTGCAAGAGACACAGCTTCAACTCTTA 300
QY 879 TGTAGAGGTGCCATTTGCTGTGAGGCGCATGCGGCTGAGATGCCCTGTCTGAGGCTGC 938
DB 301 TGTAGAGGTGCCATTTGCTGTGAGGCGCATGCGGCTGAGATGCCCTGTCTGAGGCTGC 360
QY 939 CTACCTGTCCAAAGCGGGGGCGGTGTGCTGAGAGACCTTGTGAGTCCATCCAGATGATGA 998
DB 361 CTACCTGTCCAAAGCGGGGGCGGTGTGCTGAGAGACCTTGTGAGTCCATCCAGATGATGA 420
QY 999 CCTGCTCTTACCGCTTCTCTCCAAAGGCGCAAGAGCGGAAATGAAATCCCTGAGTGA 1038
DB 421 CCTGCTCTTACCGCTTCTCTCCAAAGGCGCAAGAGCGGAAATGAAATCCCTGAGTGA 480
QY 1059 GGCCCTGTGATGCTTATCTTGAAGCAGATTAATGACGATTAAGAGCGGCTGAGTGC 1118
DB 481 GGCCCTGTGATGCTTATCTTGAAGCAGATTAATGACGATTAAGAGCGGCTGAGTGC 540
QY 1119 TTGTTACCGGGGCGAGGAGCAAGCTGAGACCTGAGCTGCTCAAGGTGAAGACATCCCTG 1178
DB 541 TTGTTACCGGGGCGAGGAGCAAGCTGAGACCTGAGCTGCTCAAGGTGAAGACATCCCTG 600
QY 1179 GAGCACTGTGCTTTTATCCATTTGACATTTGAGTGGCTGAGCATGATGCTCCCT 1238
DB 601 GAGCACTGTGCTTTTATCCATTTGACATTTGAGTGGCTGAGCATGATGCTCCCT 660
QY 1239 GAGAGTGTCCGACATGCTGTGAGATTCCTGCTTCAAGAGACAGGAGCGGATGAC 1298
DB 661 GAGAGTGTCCGACATGCTGTGAGATTCCTGCTTCAAGAGACAGGAGCGGATGAC 720
QY 1299 GTTGTTCATGCTATGCTTCAAGAGACATCTTGTGAGCTTTGAGGAGACCAAGAGTGC 1358
DB 721 GTTGTTCATGCTATGCTTCAAGAGACATCTTGTGAGCTTTGAGGAGACCAAGAGTGC 780
QY 1359 CAAGCTGAAGAGATCC 1375
DB 781 CAAGCTGAAGAGATGC 797

RESULT 9
US-10-245-851-91
Sequence 91, Application US/10245851
Publication No. US2003068782A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Bacon, Dan
APPLICANT: Filvaroff, Ellen
APPLICANT: Goddard, Audrey
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin
APPLICANT: Smith, Victoria
APPLICANT: Stephan, Jean-Phillippe
APPLICANT: Matamde, Colin
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
APPLICANT: Fong, Sherman
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3630R1C93
CURRENT APPLICATION NUMBER: US/10/245, 851
CURRENT FILING DATE: 2002-09-16
PRIOR APPLICATION NUMBER: 10/197942
PRIOR FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/059114
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/063046

PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/065027
PRIOR FILING DATE: 1997-11-10
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/086478
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090689
PRIOR FILING DATE: 1998-06-25
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 116
SEQ ID NO 91
LENGTH: 2597
TYPE: DNA
ORGANISM: Homo Sapien
US-10-245-851-91
Query Match 13.9%; Score 793.8; DB 14; Length 2597;
Best Local Similarity 99.7%; Pred. No. 1,6e-215;
Matches 795; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 579 CGAGTATTTTCCACCATCTTCCAGCCGGAAATGACCAAGAACTTGAGGCGGATGCA 638
DB 1 CGAGTATTTTCCACCATCTTCCAGCCGGAAATGACCAAGAACTTGAGGCGGATGCA 60
QY 639 GTTGGGCTAGCTTCCATGATGATGCTGAGTGGCTGATGATGATGATGATGATGATGATGAT 698
DB 61 GTTGGGCTAGCTTCCATGATGATGCTGAGTGGCTGATGATGATGATGATGATGATGATGAT 120
QY 699 CTTACCATCATCTCTGACTTGTATGATCTATGCTATGCTTTTGAAGATGCACTT 758
DB 121 CTTACCATCATCTCTGACTTGTATGATCTATGCTATGCTTTTGAAGATGCACTT 180
QY 759 TGTCTACTTTTGAAGCTTCCAAAGCTGAGATGCTGCTCCAGAGGCTCCACCAAGGA 818
DB 181 TGTCTACTTTTGAAGCTTCCAAAGCTGAGATGCTGCTCCAGAGGCTCCACCAAGGA 240
QY 819 GCAAGGTGTATACATCCAAAGCTGTGAGGCTTTGCAAGAGACACAGCTTCAACTCTTA 878
DB 241 GCAAGGTGTATACATCCAAAGCTGTGAGGCTTTGCAAGAGACACAGCTTCAACTCTTA 300
QY 879 TGTAGAGGTGCCATTTGCTGTGAGGCGCATGCGGCTGAGATGCCCTGTCTGAGGCTGC 938
DB 301 TGTAGAGGTGCCATTTGCTGTGAGGCGCATGCGGCTGAGATGCCCTGTCTGAGGCTGC 360
QY 939 CTACCTGTCCAAAGCGGGGGCGGTGTGCTGAGAGACCTTGTGAGTCCATCCAGATGATGA 998
DB 361 CTACCTGTCCAAAGCGGGGGCGGTGTGCTGAGAGACCTTGTGAGTCCATCCAGATGATGA 420
QY 999 CCTGCTCTTACCGCTTCTCTCCAAAGGCGCAAGAGCGGAAATGAAATCCCTGAGTGA 1038
DB 421 CCTGCTCTTACCGCTTCTCTCCAAAGGCGCAAGAGCGGAAATGAAATCCCTGAGTGA 480
QY 1059 GGCCCTGTGATGCTTATCTTGAAGCAGATTAATGACGATTAAGAGCGGCTGAGTGC 1118
DB 481 GGCCCTGTGATGCTTATCTTGAAGCAGATTAATGACGATTAAGAGCGGCTGAGTGC 540
QY 1119 TTGTTACCGGGGCGAGGAGCAAGCTGAGACCTGAGCTGCTCAAGGTGAAGACATCCCTG 1178
DB 541 TTGTTACCGGGGCGAGGAGCAAGCTGAGACCTGAGCTGCTCAAGGTGAAGACATCCCTG 600
QY 1179 GAGCACTGTGCTTTTATCCATTTGACATTTGAGTGGCTGAGCATGATGCTCCCT 1238
DB 601 GAGCACTGTGCTTTTATCCATTTGACATTTGAGTGGCTGAGCATGATGCTCCCT 660
QY 1239 GAGAGTGTCCGACATGCTGTGAGATTCCTGCTTCAAGAGACAGGAGCGGATGAC 1298

PRIOR APPLICATION NUMBER: 60/059114
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/063046
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/065027
PRIOR FILING DATE: 1997-11-10
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/086478
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090689
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/091358
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/099603
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/106932
PRIOR FILING DATE: 1998-11-03
PRIOR APPLICATION NUMBER: 60/115554
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119342
PRIOR FILING DATE: 1999-02-09
PRIOR APPLICATION NUMBER: 60/123957
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: 60/123972
PRIOR FILING DATE: 1999-03-11
PRIOR APPLICATION NUMBER: 60/127372
PRIOR FILING DATE: 1999-04-01
PRIOR APPLICATION NUMBER: 60/131271
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/133459
PRIOR FILING DATE: 1999-05-11
PRIOR APPLICATION NUMBER: 60/135725
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/135729
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/135750
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/138385
PRIOR FILING DATE: 1999-06-09
PRIOR APPLICATION NUMBER: 60/140653
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/141037
PRIOR FILING DATE: 1999-06-23
PRIOR APPLICATION NUMBER: 60/144732
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/144758
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/144790
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145228
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: 60/146222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: 60/146843
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: 60/148188
PRIOR FILING DATE: 1999-08-10
PRIOR APPLICATION NUMBER: 60/148513
PRIOR FILING DATE: 1999-08-12
PRIOR APPLICATION NUMBER: 60/149327
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/149395

PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/150114
PRIOR FILING DATE: 1999-08-20
PRIOR APPLICATION NUMBER: 60/151700
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 60/151734
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 60/162506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 60/170262
PRIOR FILING DATE: 1999-12-09
PRIOR APPLICATION NUMBER: 60/177118
PRIOR FILING DATE: 2000-01-20
PRIOR APPLICATION NUMBER: 60/179651
PRIOR FILING DATE: 2000-02-02
PRIOR APPLICATION NUMBER: 60/180921
PRIOR FILING DATE: 2000-02-08
PRIOR APPLICATION NUMBER: 60/187202
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: 60/198587
PRIOR FILING DATE: 2000-04-18
PRIOR APPLICATION NUMBER: 60/199614
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 60/206330
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/206368
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/209832
PRIOR FILING DATE: 2000-06-05
PRIOR APPLICATION NUMBER: 60/218371
PRIOR FILING DATE: 2000-07-13
PRIOR APPLICATION NUMBER: 60/222695
PRIOR FILING DATE: 2000-08-02
PRIOR APPLICATION NUMBER: 60/229896
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/230621
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/232887
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 60/235147
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 60/261878
PRIOR FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: 60/261910
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/261939
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/262150
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/264395
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/266421
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/267623
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/274339
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 60/280982
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: 60/282129
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/282199
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/290589
PRIOR FILING DATE: 2001-05-09
PRIOR APPLICATION NUMBER: 60/180997
PRIOR FILING DATE: 1998-11-19
PRIOR APPLICATION NUMBER: 60/267213
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: 60/380137
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 60/380138
PRIOR FILING DATE: 1999-08-25

Mon Feb 23 09:34:02 2004

us-09-964-956-12.rmpb

Page 17

PRIOR APPLICATION NUMBER: 09/403297
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 09/423741
PRIOR FILING DATE: 1999-11-10
PRIOR APPLICATION NUMBER: 09/709238
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 09/802706
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 09/872035
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 09/924419
PRIOR FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: 09/927796
PRIOR FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 09/929404
PRIOR FILING DATE: 2001-08-13
PRIOR APPLICATION NUMBER: 09/931836
PRIOR FILING DATE: 2001-08-16
PRIOR APPLICATION NUMBER: 09/941992
PRIOR FILING DATE: 2001-08-28
PRIOR APPLICATION NUMBER: 09/946374
PRIOR FILING DATE: 2001-09-04
PRIOR APPLICATION NUMBER: 10/001054
PRIOR FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: 10/052586
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 10/081056
PRIOR FILING DATE: 2002-02-20
PRIOR APPLICATION NUMBER: 10/119480
PRIOR FILING DATE: 2002-04-09

Query Match 13.9%; Score 793.8; DB 14; Length 2597;
Best Local Similarity 99.7%; Pred. No 1.6e-215; Indels 0; Gaps 0;
Matches 795; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 579 CGAGTATTTTCCCAACATCTCCAGCCGGAAGTACACCAAGTCTGAGCGGATGCAAT 638
DB 1 CGAGTATTTTCCCAACATCTCCAGCCGGAAGTACACCAAGTCTGAGCGGATGCAAT 60
QY 639 GTTCGCGTACGCTTTCCATGATGATGCTGCGCTGATGATTAAGTCCCTCGACAC 698
DB 61 GTTCGCGTACGCTTTCCATGATGATGCTGCGCTGATGATTAAGTCCCTCGACAC 120
QY 699 CTTACACATCATCCCTGATCTTGAATCTATGCTATGCTTTTGAAGTGGCAACTT 758
DB 121 CTTACACATCATCCCTGATCTTGAATCTATGCTATGCTTTTGAAGTGGCAACTT 180
QY 759 TGTCTACTTTTGAACCTTCCCACTGATGATGCTTCCCAAGGCTCCCAACCAAGGA 818
DB 181 TGTCTACTTTTGAACCTTCCCACTGATGATGCTTCCCAAGGCTCCCAACCAAGGA 240
QY 819 GAGAGTGTATACATCCCAAGTCTGAGGCTTTGCAAGAGGACACAGCTTCAACTCTTA 878
DB 241 GAGAGTGTATACATCCCAAGTCTGAGGCTTTGCAAGAGGACACAGCTTCAACTCTTA 300
QY 879 TGTNAGGTGCGCATTTGCTGTGACGCGAGTGGGGTGAAGTACCGCTGTGACAGGCTGC 938
DB 301 TGTNAGGTGCGCATTTGCTGTGACGCGAGTGGGGTGAAGTACCGCTGTGACAGGCTGC 360
QY 939 CTACTGTCCAAAGGGGGGGCGGCTTGGCAGAGACCTTGAAGTCCATCAATGATGA 998
DB 361 CTACTGTCCAAAGGGGGGGCGGCTTGGCAGAGACCTTGAAGTCCATCAATGATGA 420
QY 999 CCGTGTCTTCAACCGTCTTCTCCAAAGGCGCAAGCGGAAATGAATTCCTGATGATGTC 1058
DB 421 CCGTGTCTTCAACCGTCTTCTCCAAAGGCGCAAGCGGAAATGAATTCCTGATGATGTC 480
QY 1059 GGGCGGTGCAATCTTCAATCTTGAAGCAGATTAATGACCGCATTAAGAGCGGCTGACGTC 1118
DB 481 GGGCGGTGCAATCTTCAATCTTGAAGCAGATTAATGACCGCATTAAGAGCGGCTGACGTC 540

QY 1119 TTGTTACCGGGGGGAGGAGCAAGCTGACCTGAGCTGCAAGTGAAGACATCCCTG 1178
DB 541 TTGTTACCGGGGGGAGGAGCAAGCTGACCTGAGCTGCAAGTGAAGACATCCCTG 600
QY 1179 CAGCAGTGGCGCTCTTAACCAATGACGATTAATCTTGTGCGCTTGAATGCTCCCT 1238
DB 601 CAGCAGTGGCGCTCTTAACCAATGACGATTAATCTTGTGCGCTTGAATGCTCCCT 660
QY 1239 GGGAGTTCGCAATGGTGGTGAATTCCTGCTTACGAGAGACAGGACCGCATGAC 1298
DB 661 GGGAGTTCGCAATGGTGGTGAATTCCTGCTTACGAGAGACAGGACCGCATGAC 720
QY 1299 GTCTGTATGCAATGATGCTTCAAGAACCAACTCTTGGGCTTTGTGGGACCAAAAGTGG 1358
DB 721 GTCTGTATGCAATGATGCTTCAAGAACCAACTCTTGGGCTTTGTGGGACCAAAAGTGG 780
QY 1359 CAAGCTGAAGAAATCC 1375
DB 781 CAAGCTGAAGAAATGTC 797

RESULT 12
US-10-238-183-91
Sequence 91, Application US/10238183
Publication No. US20030073189X1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Bacon, Dan
APPLICANT: Filvaroff, Ellen
APPLICANT: Goddard, Audrey
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Guiney, Austin
APPLICANT: Smith, Victoria
APPLICANT: Stephan, Jean-Philippe
APPLICANT: Watande, Colin
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
APPLICANT: Fong, Sherman
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3630R1C11
CURRENT APPLICATION NUMBER: US/10/238, 183
CURRENT FILING DATE: 2002-09-09
PRIOR APPLICATION NUMBER: 10/197942
PRIOR FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/059114
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/063046
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/065027
PRIOR FILING DATE: 1997-11-10
PRIOR APPLICATION NUMBER: 60/073689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/086478
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090689
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/091358
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/099803
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/106932
PRIOR FILING DATE: 1998-11-03
PRIOR APPLICATION NUMBER: 60/115554
PRIOR FILING DATE: 1999-01-12

Mon Feb 23 09:34:02 2004

us-09-964-956-12.rnpb

Page 18

PRIOR APPLICATION NUMBER: 60/119342
PRIOR FILING DATE: 1999-02-09
PRIOR APPLICATION NUMBER: 60/123957
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: 60/123972
PRIOR FILING DATE: 1999-03-11
PRIOR APPLICATION NUMBER: 60/127372
PRIOR FILING DATE: 1999-04-01
PRIOR APPLICATION NUMBER: 60/131271
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/133459
PRIOR FILING DATE: 1999-05-11
PRIOR APPLICATION NUMBER: 60/135725
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/135729
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/135750
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/138385
PRIOR FILING DATE: 1999-06-09
PRIOR APPLICATION NUMBER: 60/140653
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/141037
PRIOR FILING DATE: 1999-06-23
PRIOR APPLICATION NUMBER: 60/144732
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/144758
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/144790
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145228
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: 60/146222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: 60/146843
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: 60/148188
PRIOR FILING DATE: 1999-08-10
PRIOR APPLICATION NUMBER: 60/148513
PRIOR FILING DATE: 1999-08-12
PRIOR APPLICATION NUMBER: 60/149327
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/149395
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/150114
PRIOR FILING DATE: 1999-08-20
PRIOR APPLICATION NUMBER: 60/151700
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 60/151734
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 60/162506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 60/170262
PRIOR FILING DATE: 1999-12-09
PRIOR APPLICATION NUMBER: 60/171118
PRIOR FILING DATE: 2000-01-20
PRIOR APPLICATION NUMBER: 60/179851
PRIOR FILING DATE: 2000-02-02
PRIOR APPLICATION NUMBER: 60/180921
PRIOR FILING DATE: 2000-02-08
PRIOR APPLICATION NUMBER: 60/187202
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: 60/198587
PRIOR FILING DATE: 2000-04-18
PRIOR APPLICATION NUMBER: 60/199614
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 60/206330
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/206368
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/209832

PRIOR FILING DATE: 2000-06-05
PRIOR APPLICATION NUMBER: 60/218371
PRIOR FILING DATE: 2000-07-13
PRIOR APPLICATION NUMBER: 60/222695
PRIOR FILING DATE: 2000-08-02
PRIOR APPLICATION NUMBER: 60/229896
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/230621
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/232887
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 60/235147
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 60/261878
PRIOR FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: 60/261910
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/261939
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/262150
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/264395
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/266421
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/267623
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/274399
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 60/280982
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: 60/282129
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/282199
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/290589
PRIOR FILING DATE: 2001-05-09
PRIOR APPLICATION NUMBER: 60/180997
PRIOR FILING DATE: 1998-11-19
PRIOR APPLICATION NUMBER: 60/267213
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: 60/380137
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 60/380138
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 60/403297
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 60/423741
PRIOR FILING DATE: 1999-11-10
PRIOR APPLICATION NUMBER: 60/709238
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/802706
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 60/872035
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 60/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/924419
PRIOR FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: 60/927796
PRIOR FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 60/929404
PRIOR FILING DATE: 2001-08-13
PRIOR APPLICATION NUMBER: 60/931836
PRIOR FILING DATE: 2001-08-16
PRIOR APPLICATION NUMBER: 60/941592
PRIOR FILING DATE: 2001-08-28
PRIOR APPLICATION NUMBER: 60/946374
PRIOR FILING DATE: 2001-09-04
PRIOR APPLICATION NUMBER: 60/001054
PRIOR FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: 10/052586
PRIOR FILING DATE: 2002-01-15

Page 19

QY	579	CGAGTATTTTCCCAACATCTCCAGCCGGAAACTGACCAAGAACTCTGAGGGGAGATGGCAT	638
Db	1	CGAGTATTTTCCCAACATCTCCAGCCGGAAACTGACCAAGAACTCTGAGGGGAGATGGCAT	60
QY	639	GTTGCGGTACGCTCTTCCATGATGATGTGTGGTGCCTGATGATTAAGATCCCTTCGGACAC	698
Db	61	GTTGCGGTACGCTCTTCCATGATGATGTGTGGTGCCTGATGATTAAGATCCCTTCGGACAC	120
QY	699	CTTCAACCATCATCCCTGACTTTGATATCTACTATGTCTATGTGTTTAAAGATGGGCACTT	758
Db	121	CTTCAACCATCATCCCTGACTTTGATATCTACTATGTCTATGTGTTTAAAGATGGGCACTT	180
QY	759	TGTCTACTTTTGGACCCCTCCAACTGAGATGGTGTCTCCACAGGCTCCACCAACAAGAA	818
Db	181	TGTCTACTTTTGGACCCCTCCAACTGAGATGGTGTCTCCACAGGCTCCACCAACAAGAA	240
QY	819	GCAGGTGTATACATCCAAGCTGTGATAGGCTTTTGCAAAGAGGACACAGCCTTCACTCTTA	878
Db	241	GCAGGTGTATACATCCAAGCTGTGATAGGCTTTTGCAAAGAGGACACAGCCTTCACTCTTA	300
QY	879	TGTAGAGGTGCCCATTTGGCTGTAGGCGCATGGGGTGAATATCCGCTCTCTCCAGGCTGC	938
Db	301	TGTAGAGGTGCCCATTTGGCTGTAGGCGCATGGGGTGAATATCCGCTCTCTCCAGGCTGC	360
QY	939	CTACCTGTCCAAAGCGGGGGCCGCTGTGGCAGAGACCTTGGAGTTCATCCAGATGATGA	998
Db	361	CTACCTGTCCAAAGCGGGGGCCGCTGTGGCAGAGACCTTGGAGTTCATCCAGATGATGA	420
QY	999	CCTGCTCTTCCACCGTCTTCTCCAAAGGGCCAGAGCGGAAATGAATCCCTGATGAGTTC	1058
Db	421	CCTGCTCTTCCACCGTCTTCTCCAAAGGGCCAGAGCGGAAATGAATCCCTGATGAGTTC	480
QY	1059	GGCCCTGTGCATCTTCATCTTGAAGATGAATTAATGCCCATTTAAGAACCGGCTGAGTTC	1118
Db	481	GGCCCTGTGCATCTTCATCTTGAAGATGAATTAATGCCCATTTAAGAACCGGCTGAGTTC	540
QY	1119	TGTGTTAACGGGGGAGAGGAGCGGTGACCTGGCCTTGCTCAAGGTGAAGAACATCCCTCG	1178
Db	541	TGTGTTAACGGGGGAGAGGAGCGGTGACCTGGCCTTGCTCAAGGTGAAGAACATCCCTCG	600
QY	1179	CAGCATGTGCGCTCTTAACCATTTGACGATAACTTTGTGTGCTCTGACATGATGTCTCCCT	1238
Db	601	CAGCATGTGCGCTCTTAACCATTTGACGATAACTTTGTGTGCTCTGACATGATGTCTCCCT	660
QY	1239	GGGAGTGTCCGACATGTGTGCGTGGAAATTCCTCGCTTCAAGGAGGACAGGAGCCGCAATGAC	1298
Db	661	GGGAGTGTCCGACATGTGTGCGTGGAAATTCCTCGCTTCAAGGAGGACAGGAGCCGCAATGAC	720
QY	1299	GTTCTGTTCATCGATATGTCTACAAAGAACACTCTCTGGCCTTTGTGTGGACCAAAAAGTGA	1358
Db	721	GTTCTGTTCATCGATATGTCTACAAAGAACACTCTCTGGCCTTTGTGTGGACCAAAAAGTGA	780
QY	1359	CAAGCTGAAGAAAGATCC	1375
Db	781	CAAGCTGAAGAAAGTGC	797

Query Match	13.9%	Score 793.8;	DB 14;	Length 2557;
Best Local Similarity	99.7%;	Pred. No. 1.6e-215;		
Matches 795; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

QY	579	CGAGTATTTTCCCAACCATCTCCACCGGAAAC	CTACACAGAAACCTGAGGGGAGATGGCAT	638
Db	1	CGAGTATTTTCCCAACCATCTCCACCGGAAAC	CTACACAGAAACCTGAGGGGAGATGGCAT	60
QY	639	GTTGGCGAGACGCTTCCATGATGAGTTCGTG	GGCCCTCGATGATTAAGATCCCTTGGACAC	698
Db	61	GTTGGCGAGACGCTTCCATGATGAGTTCGTG	GGCCCTCGATGATTAAGATCCCTTGGACAC	120
QY	699	CTTCAACCATCATCCCTGACTTGTATATCTAT	GTCTATGTCTATGATTTTACAGTGGCAACT	758
Db	121	CTTCAACCATCATCCCTGACTTGTATATCTAT	GTCTATGTCTATGATTTTACAGTGGCAACT	180
QY	759	TGTCACTTTTGAACCCCTCCAACTGAGATGTG	TCCTCCACAGCTCCACCAACCAAGA	818
Db	181	TGTCACTTTTGAACCCCTCCAACTGAGATGTG	TCCTCCACAGCTCCACCAACCAAGA	240
QY	819	GCAAGGTATATCATTCCAAGCTGTGAGGCTT	TGCAGAAGAGACACAGGCTTCAACTCTCTA	878
Db	241	GCAAGGTATATCATTCCAAGCTGTGAGGCTT	TGCAGAAGAGACACAGGCTTCAACTCTCTA	300
QY	879	TGTAGAGGTGCCATTGGCTGTGAGCGCAGTGG	GGGTGAGTACCGCTGTCCAGAGCTGC	938
Db	301	TGTAGAGGTGCCATTGGCTGTGAGCGCAGTGG	GGGTGAGTACCGCTGTCCAGAGCTGC	360
QY	939	CTACTGTCCAAAGCGGGGCGGTGCTTGGCAG	AGACCTTTGGAATTCATCCGATGATGA	998

Mon Feb 23 09:34:02 2004

us-09-964-956-12.rnpb

Page 20

Db	361	CTACGTGTCACAAAGCGGGGGCGGTGCTTGGCAGGACCGTTGGAGTCCATCCAGATGATGA	420
Qy	999	CCGTGCTCTTTCACCGGTCTTCTCCAAAGGCCAGAAAGCGGAAATTAATCCCTGGATGAATGC	1056
Db	421	CCGTGCTCTTTCACCGGTCTTCTCCAAAGGGCGAAGGGGAAATGAATGAATCCCTGGATGAATGC	480
Qy	1059	GGCCCTGTGATATTCTATCTTGAAAGCAGATTAATGACCGCATTAAGAGAGCGGTGTGAGTC	1116
Db	481	GGCCCTGTGATATTCTATCTTGAAAGCAGATTAATGACCGCATTAAGAGAGCGGTGTGAGTC	540
Qy	1119	TTGTATACCGGAGGCGAGGGGACGCTGGACCTGGCTGGCTCAAGGTGAAGGACATCCCTCG	1176
Db	541	TTGTATACCGGAGGCGAGGGGACGCTGGACCTGGCTGGCTCAAGGTGAAGGACATCCCTCG	600
Qy	1179	CAGAGAGCGCTCTTAAACCATGAGAGATTACTCTGTGTGGCTGGACATGAAATGCTCCCT	1236
Db	601	CAGAGAGCGCTCTTAAACCATGAGAGATTACTCTGTGTGGCTGGACATGAAATGCTCCCT	660
Qy	1239	GGGAGTGTCCGACATGATGTCGTGAATTCGCCGTCTTACGAGAGACAGGACCCGATGAC	1296
Db	661	GGGAGTGTCCGACATGATGTCGTGAATTCGCCGTCTTACGAGAGACAGGACCCGATGAC	720
Qy	1299	GTCTGTATCGCATATGTCATACAAAGAACACTCTCTGGCCTTTGTGGGACCCAAATGTG	1356
Db	721	GTCTGTATCGCATATGTCATACAAAGAACACTCTCTGGCCTTTGTGGGACCCAAATGTG	780
Qy	1359	CAAGCTGAAGAAAGATCC	1375
Db	781	CAAGCTGAAGAAAGTGC	797

```

RESULT 14
US-10-238-370-91
/ Sequence 91, Application US/10238370
/ Publication No. US20030073191A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin
/ APPLICANT: Eaton, Dan
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Grimaldi, J. Christopher
/ APPLICANT: Guiney, Austin
/ APPLICANT: Smith, Victoria
/ APPLICANT: Stephan, Jean-Philippe
/ APPLICANT: Watande, Colin
/ APPLICANT: Wood, William
/ APPLICANT: Zhang, Zemin
/ APPLICANT: Fong, Sherman
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE OF INVENTION: ACIDS ENCODING THE SAME
/ FILE REFERENCE: P3630R1C10
/ CURRENT APPLICATION NUMBER: US/10/238,370
/ CURRENT FILING DATE: 2002-09-09
/ PRIOR APPLICATION NUMBER: 10/197942
/ PRIOR FILING DATE: 2002-07-18
/ PRIOR APPLICATION NUMBER: 60/059114
/ PRIOR FILING DATE: 1997-09-17
/ PRIOR APPLICATION NUMBER: 60/063046
/ PRIOR FILING DATE: 1997-10-24
/ PRIOR APPLICATION NUMBER: 60/065027
/ PRIOR FILING DATE: 1997-11-10
/ PRIOR APPLICATION NUMBER: 60/079689
/ PRIOR FILING DATE: 1998-03-27
/ PRIOR APPLICATION NUMBER: 60/086478
/ PRIOR FILING DATE: 1998-05-22
/ PRIOR APPLICATION NUMBER: 60/087607
/ PRIOR FILING DATE: 1998-06-02
/ PRIOR APPLICATION NUMBER: 60/089801
/ PRIOR FILING DATE: 1998-06-18
/ PRIOR APPLICATION NUMBER: 60/090557
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090689
/ PRIOR FILING DATE: 1998-06-25
/

```

```

? Remaining Seq Application data removed - See File Wrapper or PALM.
? NUMBER OF SEQ ID NOS: 116
? SEQ ID NO 91
? LENGTH: 2597
? TYPE: DNA
? ORGANISM: Homo Sapien
US-10-238-370-91

Query March 13.9% Score 793.8; DB: 14; Length 2597;
BestLocal Similarity 99.7%; Pred No: 1,6e-215;
Matches 795; Conservative 2; Indels 0; Gaps 0;

```

Query Match	Similarity	99.7%	Score 793.8	DB 14	Length 2597
Best Local	Similarity	99.7%	Pred. No. 1.6e-215		
Matches 795	Conservative	0	Mismatches 2	Indels 0	Gaps 0
QY	579	CGAGTATTTTCCACCACTTCACGCGGAAA	CTGACCAAGAACTCTGAGCGGATGCGAT	638	
Db	1	CGAGTATTTTCCACCACTTCACGCGGAAA	CTGACCAAGAACTCTGAGCGGATGCGAT	60	
QY	639	GTTGGCGAAGGCTTCCATGATAGATGCGGCGCTGCATGATTAAGATGCCCTTGAGACAC	698		
Db	61	GTTGGCGAAGGCTTCCATGATAGATGCGGCGCTGCATGATTAAGATGCCCTTGAGACAC	120		
QY	699	CTTGACCAATCAATCCCTGACATTGATATCTA	CTATGCTATGTTTGAAGATGCAATCTT	758	
Db	121	CTTGACCAATCAATCCCTGACATTGATATCTA	CTATGCTATGTTTGAAGATGCAATCTT	180	
QY	759	TGCTACTTTTGAACCTCCAACTGATAGATGCTGCACACAGGCTCCACCAAGAA	818		
Db	181	TGCTACTTTTGAACCTCCAACTGATAGATGCTGCACACAGGCTCCACCAAGAA	240		
QY	819	GCAGGTGATATCATATCCAAAGCTCTGAGGCTTTGCAAGAGGACACAGCTTCAACTCTTA	878		
Db	241	GCAGGTGATATCATATCCAAAGCTCTGAGGCTTTGCAAGAGGACACAGCTTCAACTCTTA	300		
QY	879	TGTAGAGGTGCCATTTGGCTGTGAGCGGAGTGGGATGAGTACCGCTGTGACGGCTGC	938		
Db	301	TGTAGAGGTGCCATTTGGCTGTGAGCGGAGTGGGATGAGTACCGCTGTGACGGCTGC	360		
QY	939	CTACCTGTCCAAAGCGGCGGCGGCTTGTGCGAGGACCTTTGAGTCCATCCAGATGATGA	998		
Db	361	CTACCTGTCCAAAGCGGCGGCGGCTTGTGCGAGGACCTTTGAGTCCATCCAGATGATGA	420		
QY	999	CTGCTCTTTTCCACCGCTTTTCCAAAGGCGCAAGACCGGAAATGAAATCCCTGATGAGTC	1058		
Db	421	CTGCTCTTTTCCACCGCTTTTCCAAAGGCGCAAGACCGGAAATGAAATCCCTGATGAGTC	480		
QY	1059	GAGCCTGTGACATCTTCACTTTGAAGACAGATTAATGACCGCATTAAGACCGGCTGCAGTC	1118		
Db	481	GAGCCTGTGACATCTTCACTTTGAAGACAGATTAATGACCGCATTAAGACCGGCTGCAGTC	540		
QY	1119	TTGTTATCCGCGGCGCAAGGCGACGCTGACCTTGGCTGTGCAAGGTGAAGATATCCCTG	1178		
Db	541	TTGTTATCCGCGGCGCAAGGCGACGCTGACCTTGGCTGTGCAAGGTGAAGATATCCCTG	600		
QY	1179	CAGCAGTGCAGCTCTTAAACATTGACGATTAATCTTGTGCGCTGACATGAATGCTCCCT	1238		
Db	601	CAGCAGTGCAGCTCTTAAACATTGACGATTAATCTTGTGCGCTGACATGAATGCTCCCT	660		
QY	1239	GGGAGTGTCCACATGATGCTGTGATATCCCTGCTTCAACGGAAGACAGGCGCGGATGAC	1298		
Db	661	GGGAGTGTCCACATGATGCTGTGATATCCCTGCTTCAACGGAAGACAGGCGCGGATGAC	720		
QY	1299	GTCAGTATCCGCAATATGCTCAACAAGAACCACTCTGTGCGCTTTGTGAGCACCAAAAGTG	1358		
Db	721	GTCAGTATCCGCAATATGCTCAACAAGAACCACTCTGTGCGCTTTGTGAGCACCAAAAGTG	780		
QY	1359	CAAGCTGAAGAGATCC	1375		
Db	781	CAAGCTGAAGAGATGTC	797		

Mon Feb 23 09:34:02 2004

us-09-964-956-12.inpb

Page 21

GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Eaton, Dan
APPLICANT: Filvaroff, Ellen
APPLICANT: Goddard, Audrey
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Guirney, Austin
APPLICANT: Smith, Victoria
APPLICANT: Stephan, Jean-Philippe
APPLICANT: Watanabe, Colin
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
APPLICANT: Fong, Sherman
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3630R1C88
CURRENT APPLICATION NUMBER: US/10/245,055
PRIOR FILING DATE: 2002-09-16
PRIOR APPLICATION NUMBER: 10/197942
PRIOR FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/059114
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/063046
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/065027
PRIOR FILING DATE: 1997-11-10
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/086478
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090689
PRIOR FILING DATE: 1998-06-25
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 116
SEQ ID NO 91
LENGTH: 2597
TYPE: DNA
ORGANISM: Homo Sapien
US-10-245-055-91

Query Match 13.9%; Score 793.8; DB 14; Length 2597;
Best Local Similarity 99.7%; Pred. No. 1.6e-215;
Matches 795; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 579 CGAGTATTTTCCACCATCTCCAGCCGGAACCTGACCAAGAACTCTGAGGCGGATGCGAT 638
DB 1 CGAGTATTTTCCACCATCTCCAGCCGGAACCTGACCAAGAACTCTGAGGCGGATGCGAT 60
QY 639 GTTCGGGTAGCTTTCCATGATGAGTTCGTGGCTCGATGATTAAGATCCCTTGGAGAC 698
DB 61 GTTCGGGTAGCTTTCCATGATGAGTTCGTGGCTCGATGATTAAGATCCCTTGGAGAC 120
QY 699 CTTACCATCATCCCTGACTTGTATCTACTATGTCATGTTAGTGGTGGAGTGGCACTT 758
DB 121 CTTACCATCATCCCTGACTTGTATCTACTATGTCATGTTAGTGGTGGAGTGGCACTT 180
QY 759 TGTCTACTTTTGAACCTTCCAACTGAGATGTTCTCCACCAAGGCTCCACCAAGGA 818
DB 181 TGTCTACTTTTGAACCTTCCAACTGAGATGTTCTCCACCAAGGCTCCACCAAGGA 240
QY 819 GCAGGTGTATACATCCAGCTGAGGCTTTGCAAGAGGACACAGCCTTCACTCTTA 878
DB 241 GCAGGTGTATACATCCAGCTGAGGCTTTGCAAGAGGACACAGCCTTCACTCTTA 300
QY 879 TGTAGAGGTGCGCATTTGGCTGTGAGGCGAGTGGGGTGGAGTACCGCTTCTGAGGCTGC 938
DB 301 TGTAGAGGTGCGCATTTGGCTGTGAGGCGAGTGGGGTGGAGTACCGCTTCTGAGGCTGC 360

QY 939 CTACCTGTCCAAAGCCGGGGCCGTGCTTGGCAGAGACCTTTGAGTCCATCCAGATATGA 998
DB 361 CTACCTGTCCAAAGCCGGGGCCGTGCTTGGCAGAGACCTTTGAGTCCATCCAGATATGA 420
QY 999 CCGTCTTTACCGTCTTCTCCAAAGGCGCAAGAGCCGAAATGAAATCCCTGGATGATC 1058
DB 421 CCGTCTTTACCGTCTTCTCCAAAGGCGCAAGAGCCGAAATGAAATCCCTGGATGATC 480
QY 1059 GGCCCTGTGCATCTTCACTTGAAGCAATAAATGACCCCATTTAGAGAGCGGTGCAATC 1118
DB 481 GGCCCTGTGCATCTTCACTTGAAGCAATAAATGACCCCATTTAGAGAGCGGTGCAATC 540
QY 1119 TTGTTACCGGGGGGAGAGGAGCAGCTGGACCTGGCTCGATCAAGGTGAAGACATCCCTG 1178
DB 541 TTGTTACCGGGGGGAGAGGAGCAGCTGGACCTGGCTCGATCAAGGTGAAGACATCCCTG 600
QY 1179 CAGCAGTGCAGCTCTTAACCATTAAGCATTAATTCTGTGGCTTGGACATGAATGCTCCCT 1238
DB 601 CAGCAGTGCAGCTCTTAACCATTAAGCATTAATTCTGTGGCTTGGACATGAATGCTCCCT 660
QY 1239 GGGAGTGTCCGACATGTTGTGGGATTTCCCTTCAAGGAGACAGGAGCCGATGAC 1298
DB 661 GGGAGTGTCCGACATGTTGTGGGATTTCCCTTCAAGGAGACAGGAGCCGATGAC 720
QY 1299 GTCTGTATCGCATATGTTCTAACAAGACATCTCTGGCTTTGTGGGACCCAAAGTGG 1358
DB 721 GTCTGTATCGCATATGTTCTAACAAGACATCTCTGGCTTTGTGGGACCCAAAGTGG 780
QY 1359 CAAGCTGAAGAAATTC 1375
DB 781 CAAGCTGAAGAAATTC 797

Search completed: February 20, 2004, 02:47:35
Job time : 1250 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 16:15:27 ; Search time 8735 Seconds
(without alignments)
19455.715 Million cell updates/sec

Title: US-09-964-956-12

Perfect score: 5691
Sequence: 1 atgaagccatgcctcgaa.....gcttagacagctgaataaa 5691

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estdb:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hrc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hrc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrt:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2239.6	39.4	5691	29	AY420219 Homo sapi
2	2175.4	38.2	3361	11	AK051614 Mus muscu
3	2111.4	37.1	5564	29	AY420221 Mus muscu
4	2017.2	35.4	5520	29	AY420666 Homo sapi

Result No.	Score	Query Match	Length	DB ID	Description
5	1767.2	31.1	5447	29	AY420668 Mus muscu
6	1681.4	29.5	5520	29	AY420667 Pan trogl
7	1552.6	27.3	5341	29	AY420220 Pan trogl
8	1221.2	21.5	3355	11	AK049319 Mus muscu
9	743.8	13.1	838	14	CD653539 AGENCOURT
10	658.8	11.6	1059	12	BM547417 AGENCOURT
11	650.6	11.4	1201	9	AL529477
12	638.4	11.2	745	14	CA316867
13	613	10.8	618	13	EX470595
14	599.2	10.5	1027	12	BM544169
15	578	10.2	882	13	BU149880
16	556.4	9.8	646	14	CD352102
17	550.4	9.7	862	14	CB724079
18	546.4	9.6	924	10	BF312056
19	538.8	9.5	638	12	BC066250
20	525.4	9.2	831	14	CD349170
21	525.2	9.2	932	13	BU121882
22	517.6	9.1	822	14	CD802878
23	514.2	9.0	821	14	CD802734
24	513.8	9.0	867	13	BU117252
25	512.8	9.0	1037	13	BU172674
26	510.8	9.0	762	13	BU268254
27	505.4	8.9	612	10	BB656076
28	502.4	8.8	760	14	CF533779
29	501.4	8.8	844	13	BQ434703
30	500.2	8.8	798	14	CA752328
31	499.8	8.8	733	14	CD351785
32	498.8	8.8	864	13	BX720011
33	498.6	8.8	657	13	BX670344
34	495.4	8.7	762	9	AL134739
35	494.6	8.7	653	9	BE294546
36	493.8	8.7	740	14	CA328914
37	493.8	8.7	976	13	BU250116
38	492.8	8.7	725	12	BM950962
39	491.6	8.6	711	14	CB519051
40	491.6	8.6	838	12	BT824913
41	488.4	8.6	716	14	CB249549
42	487.4	8.6	789	13	BX320194
43	486.4	8.5	702	12	BC825770
44	485.8	8.5	817	14	CB525253
45	484.6	8.5	894	14	CA488958

ALIGNMENTS

RESULT 1
AY420219
LOCUS
DEFINITION Homo sapiens HCM7151 gene, VIRTUAL TRANSCRIPT, partial sequence.
ACCESSION AY420219
VERSION AY420219.1 GI:39776176
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
1 (bases 1 to 5691)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission

Mon Feb 23 09:34:03 2004

us-09-964-956-12.rst

Page 2

JOURNAL		Submitted (16-Nov-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA	
COMMENT		This sequence as made by sequencing genomic exons and ordering them based on alignment.	
FEATURES		Location/Qualifiers	
source		1..5691	
gene		/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
		<1..>5691	
		/locus_tag="HOM7151"	
ORIGIN			
Query Match	39.4%	Score 2239.6	DB 29; Length 5691;
Best Local Similarity	58.0%	Pred. No. 0;	
Matches 3228;	Conservative	0; Mismatches 2299;	Indels 42; Gaps 8;
141	CGCGGAGGATTCAATCACTCTGCTGTGTGATGAGAGAGACAGACATTTACTTGGGGGC	200	
Db	141	CGACTGGGGGCGCTCACCCCACTAGTGTGATGAGAGACAGCGAGGTGTATGTGGGGCG	200
QY	201	CGTCATGGGATTATCAAGCTCTCCAGGCACTGAAAGCTTGTGTAGCGATGAGACAG	260
Db	201	AGTGAACCGCATCTATTAAGCTGTGGGAACTGACATCTGCTGGGGGCCACGTACCGG	260
QY	261	GCCGAGCAGAGACACCCCAAGTGTATCCACCCCGCATGTGTCCAGACTGATGAGCC	320
Db	261	CCCTGTGAGAGGACACAGAAATGTCTACCCCGCCGACGTCGAGTCTGTGCCCCACG	320
QY	321	CCTGACCCACCAACCAATGTCAACAAGATGTCTCTCATAGCTACAGAGAACAGGCT	380
Db	321	CTTGGGCACTATGTACACGTATACAGATGTCTGTCTGTGACTATGCCGCTAACCGCT	380
QY	381	GATTGCTGTGGAGCCTGTACCAAGCATCTGCAAGCTGTGAGGCTGAGACCTCTT	440
Db	381	GCTGGCTGTGGCAGCGCTCCCAAGGCATCTGCCAGTTCTGCGCTGTGAGCATCTCTT	440
QY	441	CAAGCTGGGGAGCCTTATCATTAAGAGAGCATATCTGTGAGTGTCAAGAGAGCG	500
Db	441	CAAACTGGGTGACCAACACACCGTAAAGGACCTAATCTGTCAAGGTGACAGAGGACG	500
QY	501	CTCAGTCTTTGGAGTATGCTCTCTAC-----AGCAACTGTGATGACAAAGCTGTTCAT	554
Db	501	CAGCATGGCGGGCGTGTCTATTTGCGGGGCAACGGGCGCAGGCGCCAAACCTCTTCAT	560
QY	555	TGCCACGGAGTGTGATGGGAAGCCGAGATATTTTCCACATCTCCACCGGAAACTGAC	614
Db	561	GGGACACACCATTCATGATGGCAAGTCCGAGTCTCCCAACATGTCACGCGGTGTCTAT	620
QY	615	CAGAACCTTGAGGCGGATGAGCATGTTCGCGTACGCTTCCATGATGATGTGTGCGCTC	674
Db	621	GGCCAAACGAGAGGATGCGGCAATGTTGGGCTTCTGTATCAAGATGATTTGTGTCTATC	680
QY	675	GATGATTAAGATCCCTTGGGACACTTTCACATCATCTCTGACTTTGATATCTACTATCT	734
Db	681	ACAGCTCAAGATCCCTTGGGACAGCTGTCTCAAGTTCCGGGCTTTGACATCTACTATCT	740
QY	735	CTATGGTTTATGAGTGGCAACTTTGTCTACTTTTGTGACCTTCCAACTGAGATGTGTCTC	794
Db	741	GTACAGCTTCCGACGAGACAGATTTGTCTACTACTCAAGCTGACGTAGACACAGCT	800
QY	795	TCCACCAAGCTCCACCAACGAGAGCAGGTGTATACATCCAAAGCTGTGAGCTTTGCCAA	854
Db	801	GACTTCGCTGTATGCGCGCGGAGACATCTTTCAGGTCCAAAGATGTGCGGTCTGTCT	860
QY	855	GGAGGACACAGCTTCAACTCTCATATAGAGTGCCATATGGCTGTGACCGGAGTGGGT	914
Db	861	GGAACACCCCAAAATTCTACTGTGATAGTTCCCATTTGGGTGAGAGCGCGGTGT	920
QY	915	GGAGTACCGGCTGTCTGCAAGCTGTCTTACTCTGTCCAAAGGGGGGCGGTCTTGGACGAC	974
Db	921	GGAGTACCGGCTGTCTGCAAGATCTCTTACTGACCGGCGCGGTGTGCTGTGCCACCA	980

QY	973	CCTTGGAGTCCATCCAGATGATGACTGCTCTTCAACCGCTTCCAAAGGCCAAGAGG	1034
Db	981	GCTGGGCGCTGGAGGACGAGACGTGCTGTTCACTGTGTTGCGCCAGGCCAAGAA	1040
QY	1035	GAAATGAATCCCTGGATGAGTGGCGCCGTGTGCATCTTCATCTTGAACAGATAATGA	1094
Db	1041	CCGGTGAACCCAAAGAGTCACACACTGTGTGCTGTTACAGCTCAGGGCATCAAGGA	1100
QY	1095	CCGCATTAAAGAACGGCTGAGTCTTGTTAACCGGGCGAGGGGACGCTGGACTGGCCGG	1154
Db	1101	GAAGATTAAAGACCGCATCAGTCTGTCTACCGTGTGAGGGCAGACTCTCCTGCGGTG	1160
QY	1155	GCTCAAGGTGAAGACATCCCTGACAGAGTGGCTTTAACATTGACGATTACTTGTG	1214
Db	1161	GCTGTCAACAAAGAGGTGGGCTGTCAACTGNNNNNNNNNNNNNNNNNNNNNNNNNN	1220
QY	1215	TGGCTGTGACATGAATGCTCCCTGGGAGTGTCCGACATGTGTCCGTGAATTCGCTCT	1274
Db	1221	NN	1280
QY	1275	CACGAGAGCAGGAGCCGCAATGACTGTGATGCGATATGTTCAAGAACCACTCTCT	1334
Db	1281	NN	1340
QY	1335	GGCTTTGTGGGCAACAAAGTGGCAAGCTGAAGAGATCGGGTGGATGAGCCCA---	1390
Db	1341	NN	1400
QY	1391	-----GGGGCACGCGCTCCAGATGAGACGTTGCAAGTGTGACCCCGGCCAGTCTT	1445
Db	1401	CGGTGGCGGCGCTGGCTGGCTCAAGAGCGTCTGTGCCCAAGAGGGAGGCCCATCTT	1460
QY	1446	CCGGGATATGGCTCTTCCAAAGACCAACGCAACTTACATCATGTCAAGAGGACGCT	1505
Db	1461	GCGAGACCTGATCTCAGCCCCCAACCAACGATCTTACGCCATGACGAGAGACAGT	1520
QY	1506	CACCAAGTCCCTGTGAGTCTGTGTCAATACAGAGCTGGCGGAGTGTCTTGGCTC	1565
Db	1521	GACCGGGATCCTGTGAGAGCTGTGTGCAGTACACCTCTGTGAGGTGTGTGTGGGCTC	1580
QY	1566	AGCGAGCCCACTGTGAGTGTGTGTGCACAACACTGACCCGGAAGAGCGGCTG	1625
Db	1581	ACGGGACCCCACTGT	1640
QY	1626	TGACCGGTCCAAAGAGCCCCGCAAGTGTGCTTGGAGATGAAGCAGTGTCCGGCTGAC	1685
Db	1641	TGACCGAGACAGACAGAGCCCAAGGCTTGTGTGGGACCTGTCTCAGTGTGTGACGCTGAC	1700
QY	1686	GGTGCATCCCAACAATATCTCCGCTCTCAGTACAAAGTGTCTGTGCTCTGTGAGACGTA	1745
Db	1701	TGTGTACGCCCGCAATGT	1760
QY	1746	CAATGTCCCGAGCTGTGAGTGTGCGTCAACTGTCACTTTGAGAGACTGTTCAGAGATGA	1805
Db	1761	GAAGTGTGTGACTCTTCAAGTGTGCGCAACTGTCTCTTGTGAGACTTCAACGAAATCTGA	1820
QY	1806	TGGCTGT	1865
Db	1821	GAGGTCTGTGAAGATGGCGGAACTCACTGCGCTCACCTCCGCGCGGAGGTGTGGCGC	1880
QY	1866	GATCATCAGAGAAATGGGACCAACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1925
Db	1881	CATCAACGCGGGCCAGAGAGACAGCGGGGTGTAACTTACTTAAGTTCAGAGAGAC	1940
QY	1926	CGGATGACTCTCGCCAGACCAAGCTTGTCTTCTTACAAATTCAGCGGTTCACAAATTCGTG	1985
Db	1941	AGGGAAGAAATTTGCTGT	2000
QY	1986	CCTGTGCTGT	2045
Db	2001	CCTGTGCTGT	2060
QY	2046	CCATGACCCCAAGACTGTCTTCCAGAAAGCCGAGTGAAGCTGCCCGAGACTGCC	2105

Db 2061 ACACACGTGCTGACCTGGCCCTTCCGAGAGGCGCGGTCAACGTGTGAGAGATGCGCC 2120
Qy 2106 CCAGCTGCTCGAGTGGACAAAGATCTGTGTGCCCCGTGAGAGTATCAACCTATACGCT 2165
Db 2121 ACAGATCTCTGCGCTTCCACGAGATCTACGTGCGAGTGGATGAACCCATCAACCT 2180
Qy 2166 GAAGGCCAAGAACTCCCGCAGCCCAAGTGTGGGCAAGCTGAGTACGAATGCACTCTCA 2225
Db 2181 GCGCGACAGGATCTGCTCAAGCCACAGTCAAGGCGAGCTGATAGAGGCTCTTCA 2240
Qy 2226 CATTACAGGCAAGCAGAGAGTGCCTGCGCTTGTGCTTCAACAGCTCCAGGCTACAGT 2285
Db 2241 CATCCCGGAGCGCCGCGCTGTCAACGCTGCGCTTCAACAGCTCCAGGCTGAGT 2300
Qy 2286 CCAAGAACCTCTTATCTATGAGGAGTGAATCAACAACTGCGGTGAGTTGAC 2345
Db 2301 CCAGATCTCTCGANN 2360
Qy 2346 AGTGTGTGGAATGGGCACTTCAACATTGACAAACCACTCAGATPAAGTTCACTCTA 2405
Db 2361 NNN 2420
Qy 2406 CAAGTGTGAGCCATGCGTGAAGCTGCGGGCTGTGCTCAAGGCTGACCCAGACTTCCG 2465
Db 2421 NNN 2480
Qy 2466 ATGTGCTGTGTGCGAGGCGCCAGTGCACCTGCGCAGACACTGCGCCAGCA 2525
Db 2481 NNN 2540
Qy 2526 GAGCAGTGTGTGAGT 2585
Db 2541 NNN 2600
Qy 2586 GATTAAT---CCCGGTGACAGGCCCCCGGAGAGGGGGGACCAAGTCACTATCCAGGGGA 2642
Db 2601 NNNNNCTGTCCCGGAGACGGGCCCGAGGCGAGGGGGGACGGGCTCATATCAAGGCCA 2660
Qy 2643 GAACCTGGGCTTGAATTTGCGACATGCGCTCCCATGTCAAGTTGTGTGGCTGAGT 2702
Db 2661 GAACCTGGGCTTGAATTTGCAAGACGTGCGCTGTGGGGTGGCGGTGGCAAGTGTGTG 2720
Qy 2703 CAGCCCTTGTAGTGAATGTTACATCCTGTGACAAAGATGTGTGTGTGTGTGTGTGTGT 2762
Db 2721 CAGCCCTGTGAGAGCGAGTACATCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2780
Qy 2763 CAAGCCC--AGCCAGATGACAGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2819
Db 2781 CAGCTCGGT 2840
Qy 2820 ATTATGAGCCCGGT 2879
Db 2841 CTACCGGCTGT 2900
Qy 2880 GCCCAGCCGGGGGCCCATGTCCGAGGGGACCCAAAGTACATCAAGGCAACCACTTGA 2939
Db 2901 CCCCCTCGGTGGGCTCTGTCAAGGGGCACTGTGATGTGATGAGGGAAGCAACCTGAA 2960
Qy 2940 TGCCGGAAGCACTGT 2999
Db 2961 CGCAGGCAAGTATGT 3020
Qy 3000 TCCATCTCTACATTTGT 3059
Db 3021 CTCCCGT 3080
Qy 3060 GGTGAGGT 3116
Db 3081 CATCAACATCAACGGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3140
Qy 3117 CCCCACATCGT 3176

Db 3141 CCCCACATCTGTGAGATTCGACCCGAGTGTGAGATCAACAGNNNNNNNNNNNNNNNNNN 3200
Qy 3177 CGTATGGGGGACCCACTGTGACTCATACAGAACCCCGAGTCCGTCCAGCATGAGG 3236
Db 3201 NNN 3260
Qy 3237 GAAGAGCACTCAATATCTGTGAGTTGTGAAGCTACTGAGTGTGACTGTCAAGGGCC 3296
Db 3261 NNN 3320
Qy 3297 CGCCGTGCTGTGGGTCTGTGACACAGTGTGACCTGACCGAGAGGCGCGAGGATTGG 3356
Db 3321 GTGTGTGCAACCTGT 3380
Qy 3357 CTTCACTGTGACCAAGT 3416
Db 3381 CTTGTGTGTGACCAAGT 3440
Qy 3417 TCCCAACCGGT 3476
Db 3441 CCTGACCCGTACTGT 3500
Qy 3477 GCCCATCATCTTAAAGGCAAGAACTGTATCCGCTGTGTGTGTGTGTGTGTGTGTGT 3536
Db 3501 CCACTCATCTCAAGGGCGGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3557
Qy 3537 GAACATCATCTGT 3596
Db 3558 CAACATCAAGGT 3617
Qy 3597 GCTGT 3656
Db 3618 GGT 3677
Qy 3657 GAGATATCTCCCCGGGAGT 3716
Db 3678 CGAGTTCTGT 3737
Qy 3717 CGTGACATGT 3776
Db 3737 TGT 3797
Qy 3777 CTATTAACGCAATGT 3836
Db 3798 CTACAGGCGCAAGT 3857
Qy 3837 CAACCTGTGAGT 3896
Db 3858 CAACCTGTGAGT 3917
Qy 3897 CATTCATGAGT 3956
Db 3918 CATTCATGAGT 3977
Qy 3957 TTACACATGT 4016
Db 3978 ATATGT 4037
Qy 4017 GGT 4076
Db 4038 GGT-----GAGGCGCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4085
Qy 4077 CAACAAAGAGT 4136
Db 4086 GACCAAGAGCACTTGT 4145
Qy 4137 CATGT 4196
Db 4146 CATGT 4205
Qy 4197 GTATGT 4256
Db 4206 ATATGT 4265

QY	425	TAAGAACCAACCCCAAGCTGCGGCTCAGAGAGACTGAGTCACTAGTGGCTAGAAAGTGTGAC	4316
Db	4266	CAGAGACCAACCCCAAGCTGCTACTCGCGCCGAGCTGAGTCTGGTGGCAGAAAGATGTCAAC	4325
QY	4317	CAATTGGTTACTTTCTCTCTCTACAAAGTTCTTCAGAGATGTCTGGGGAGCCCTCTT	4376
Db	4326	TAACCTGGTTACCTCTCTCTGTATATAGTTCTTCAGAGATGTGGCTGGGGAGCGCTGTT	4385
QY	4377	CTCCCTCTTCTGNTGCCATCAAGCAGCAGATGAGAGAGGGCCCATTTGACCCCATCAACGG	4436
Db	4386	CATGCTGTACTGGGCCATCAAGCAGCAGATGAGAGAGGGCCCATTTGACCCCATCAACGG	4445
QY	4437	CGAGGCCCGCTACTCTCTTGAGCGAGAGCAAGCTCATCCGCCAGCATTTGACTTCAAAAC	4496
Db	4446	TGAGGCAACGCTACTCCTTGAGTGGAGCAAGCTCATCCGCAGCAGATTTGACTTCAAGAC	4505
QY	4497	CCTGGTCTGTGCTGTGTGTCAGGCCAGACAATGGCAACAGCCCGAGGTCCCACTAAAGAT	4556
Db	4506	ACTGACCTGTACCTGTGTGAACTCTGAGATGAGAAATGCACCTGAGGTGGCGGTGAAGGG	4565
QY	4557	CCTCACTGTGACCACTACTCAGTGTGAAGAGAGATTTGATGTCCATCTTCAAGAA	4616
Db	4566	GCTGGACTGTGACACGGGTGACCCAGGCCAAGAGAACTCTGACCTCTCTTCAAGGG	4625
QY	4617	TGTGCTTGTGCTCCACGCGCCAAAGCTGCAGATATGATCTGGAGTGGCGACAGAGAG	4676
Db	4626	CGTCCCTTANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNAATGTGGCGCAGGCGG	4685
QY	4677	TGGGGCAAGGATGATCTTTGCAGGATGAAGACATACCAACAAGTTGAGATGATTTGGA	4736
Db	4686	CATGGCCGCGCTCATCTCTGCAGGACGAGGAGTCAACACAGATTAACAACGATTTGAA	4745
QY	4737	GCGACTGAACACACTGGGCCCACTACAGAGTGCACAGATGATGCTCCGTGTGCTATTAATGTC	4796
Db	4746	GAGCTGAACACACTGTGCTCATCTACAGTGAACAGAGGGTCTCTGGGTGGACGTGTGCC	4805
QY	4797	CAAGCAGGTGACAGCTTATAAGCGATGTAACTATCCACCGTCTTCAGAGCTTACAGAG	4856
Db	4806	CACACAGACGTGCGGCTTACCAATCTCCMACTCTTCACTTCAACAAAG--TCCCTCAG	4862
QY	4857	TAAATATAAATAACATGATCCCGGTACACGGGACGCCGACACGCTCCGCTCACGGAACCC	4916
Db	4863	CAGATACGAGAGCATCTGTGCCACGGCCAGACGCCCTCGACGCTGTGGCAGCGCC	4922
QY	4917	TATGATCACTCTGACCTCTGAGAGTGTGAGTCAAGATGTGGGACCTATGTGAAGAACACGA	4976
Db	4923	CATGATCACGCCCACTCGAGAGGGGACCAAGCTGTGGACCTGTGTGAAGAACACGA	4982
QY	4977	GCACGAGACAGAAAGAGGGGGGACCGGGGAGAGCAAGTGTGTGTAATCTTACCTGAC	5036
Db	4983	CCACTGGACCGAGGTGAGGGGTACCGGGCAGCAAGTGTGTGAGAGATCTACTTGAC	5042
QY	5037	CCGACTCTCTGGCCACTAAGGGGACACTGTGCAAAATTTGTGATGATACCTCTTGAACCAT	5096
Db	5043	AACGGTACTGTGGCAACAAAGGGGACACTGCAAAATTTGTGACGACCTGTGTGAACCAT	5102
QY	5097	CTTCGACAGGCAACACGTGTGCTCGCCCTGGCCCTGGCCATCAAGTACATGTTTGACTT	5156
Db	5103	CTTGACGACGGGACACCGGGGTACGCCCTGGCCCTGGCCATCAAGTACATGTTTGACTT	5162
QY	5157	CCTGATAGCAGGCTGTGATTAACATGGAATTCATGACCCGACGTCGGCCATCTGTGA	5216
Db	5163	CCTGATAGCAGGCGGCAACAGCAACAGTTCACGATGCTGACGTCGGCCCACTGTGA	5222
QY	5217	GAGCAATGTGCTGCCCTGAGGTTTGTGGGTCAACAATGATCAAGAACCCGACGTTTGTGT	5276
Db	5223	GAGCAATGTGCTGCCCTGAGGTTTGTGGGTGAACTGATCAAGAACCCACGTTTGTGT	5282
QY	5277	TGACATCAATAGAACAGATACAGAGGCGCGCTCTCTGTGGTGGCTCAGACCTTAT	5336
Db	5283	CGACATTCACAAAGACAGATACGAGAGCGCTCTTGTGGTGGTGGCTCAGACCTTAT	5342

[illegible]

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

REFERENCE 6 (bases 1 to 3361)

AUTHORS Aach, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hagiwara, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishii, K., Nomura, K., Numata, R., Ohno, M., Ohse, N., Okazaki, Y., Saito, R., Saitoh, K., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE Direct Submission (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

PLEASE VISIT OUR WEB SITE FOR FURTHER DETAILS.

URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/

Location/Qualifiers

FEATURES

1. 3361

source
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6J"
/db_xref="PANTOM.DB:D130060F15"
/db_xref="WGI:2419923"
/db_xref="taxon:10090"
/clone="D130060F15"
/cissue_type="spinal ganglion"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="12 days embryo"
/note="unlabeled protein product; putative similar to PLEXIN 2 [Mus musculus] (GPR170207, evidence: FAMY, 78.4% ID, 45.1% length, match=2553)"
/codon_start=1
/protein_id="BAC34692.1"
/db_xref="GI:26342060"
/translation="VRIPEWMSISGNTPIAVWGTHTDILIONPIKAGKHEINIC EVLNATENTCCAPALALGPHQSDLTERRPERGILINNVSLINTNTTYNPNPE EAFSPGILELPGPIILKGNLIPFAGNVAKNTLVNGEPCITVTSIDVQLCE SPNIGRHKVAVGMEYSPGMVIAVDSPLPALVAVAGGLITIVALLK RKRSDELTKRLQMDNLESEVALECKEFAELQTDHLSLSDGAGIPLDR TYKRVLPGLIEDHVLADLEPGVROGRVKGKLPALQINNSNPKLIRLESVA SEPMRDNVAVSLINTVOSLEKAVTDVLEKQIADLIDKLNNSNPKLIRLESVA EKLTNMTPIILYKKEKAGEPLSLFCALKKOEMGPDATGEARYSLSDEKLR COLEKTVLVSCEPDNVNPEPVKINCTTOVKEKIDAEKVPQSHPKRAK MDLEWQSGARMILQDEDTTKIBNKRKNTVAHQVDPGSVALVSKQVRYAAE NNSVTSRTSASRKNEMIRYTSPPSLSRRTWITPDSGVKHLVKNHSGQKSG DRGKQVSEIYILRLATKTLQKRFVDLFEFTISLARSGALPLAKWFDLEGA DKHGIDHNVHVTWNSCLPLRFVNMINKQFVFDIHKSIDACISVAQFTMDSG STSERLKGDSPNKLIYAKDIPIKMYVERVYDIDGMPSIDQDNAAVLAESRKH MNEFVMALEISFISVYKSEIIGPLHDHDCQCKGLAYKEQVITLMSDS"

ORIGIN

Query Match 38.2%; Score 2175.4; DB 11; Length 3361;
Best local similarity 90.6%; Pred. No. 0;
Matches 2320; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

3127 GTGGGATTTGACCGAATGAGCATTTGTCAGTGAACACACCCATCGCCGATGAGGGG 3186

Db 1 GTGGGATTTGACCGAATGAGCATTTGTCAGTGAACACACCCATCGCCGATGAGGGG 60
QY 3187 ACCGACCTGACCTCATACAGACCCCGAGATCCGTGCAAGCATGAGGAGAGAC 3246
Db 61 ATTCACCTGACCTCATACAGACCCCGAGATCCGTGCAAGCATGAGGAGAGAC 120
QY 3247 ATCAATATCTGAGAGTTCTGAAACGCTACTGAGATGACCTGACGCGCCCTGCT 3306
Db 121 ATCAATATCTGAGAGTTCTGAAATCTCAAGAGATGACCTGACGCGCCCTGCT 180
QY 3307 CTGGGTCTGACCACTGACCACTGACCACTGACCACTGACCACTGACCACTGAC 3366
Db 181 CTGGGTCTGACCACTGACCACTGACCACTGACCACTGACCACTGACCACTGAC 240
QY 3367 GACACGTCAGTCCCTGCTCATCTCTCAACAGCAACCACTTCACTTCACTTCA 3426
Db 241 GACACGTCAGTCCCTGCTCATCTCTCAACAGCAACCACTTCACTTCACTTCA 300
QY 3427 GTGTTTGAAGCTTTGTTGCTTCAAGAACTCTGAGCTCAAGCTGAGCCCTCATC 3486
Db 301 GTGTTTGAAGCTTTGTTGCTTCAAGAACTCTGAGCTCAAGCTGAGCCCTCATC 360
QY 3487 CTAAAGGCGAAGACCTGATCCGCTGTGCTGAGGCGCAAGTGAAGCTGACTAC 3546
Db 361 CTAAAGGCGAAGACCTGATCCGCTGTGCTGAGGCGCAAGTGAAGCTGACTAC 420
QY 3547 GTGCTGTGAGGAGAGAGCGCTGACCGTACCGTGTGATGTCAAGTGTCAAG 3606
Db 421 GTGCTGTGAGGAGAGAGCGCTGACCGTGTGATGTCAAGTGTCAAGTGTCAAG 480
QY 3607 TCCGCCAATCTATGCGGAGAGCAAGAGTATGCGCGTGTGAGTGTGAGTATCC 3666
Db 481 TCCGCCAATCTATGCGGAGAGCAAGAGTATGCGCGTGTGAGTGTGAGTATCC 540
QY 3667 CCGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3726
Db 541 CCGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
QY 3727 GCAAGTGTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3786
Db 601 GCGGTGTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 3787 AAGTCCCGGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3846
Db 661 AATCCCGGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 720
QY 3847 TCCCGTGTGAGGCTGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3906
Db 721 TCCCGTGTGAGGCTGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 3907 CTGACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3966
Db 781 CTGACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 840
QY 3967 CCGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4026
Db 841 CCGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
QY 4027 TACCGAGAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4086
Db 901 TACCGAGAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
QY 4087 GTGTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4146
Db 961 GTGTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
QY 4147 CGGAGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4206
Db 1021 CGGAGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1080
QY 4207 GATGTGCTGAAGAGTGTGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4266

Db 1081 GACGTGTGAGAGACAGCTACTGAGCTGACCTCATTTGATTAAGAACTTGAGAGACAGAACACC 1140

QY 4267 CCTAAGCTGTCTGTCTAGAGAGAGCTAGTCAAGTGTGTGAGAGATGCTGACCAATTTGTTT 4326

Db 1141 CCAAGAGCTGTCTGTCTAGAGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200

QY 4327 ACTTTCCTCTCTCTAAGAGTTCTCTCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4386

Db 1201 ACTTTCCTCTCTCTCTAAGAGTTCTCTCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260

QY 4387 TGTGCTCATAGAGAGAGATGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4446

Db 1261 TGTGCTCATAGAGAGAGATGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320

QY 4447 TACTCTTGTAG 4506

Db 1321 TACTCTTGTAG 1380

QY 4507 AGCTGTGTCAAG 4566

Db 1381 AGCTGTGTCAAG 1440

QY 4567 GACACCATCATCTAGGTCTAAAG 4626

Db 1441 GACACCATCATCTAGGTCTAAAG 1500

QY 4627 TCCACACGAG 4686

Db 1501 TCCACACGAG 1560

QY 4687 ATGATCTTGAG 4746

Db 1561 ATGATCTTGAG 1620

QY 4747 ACACTGAG 4806

Db 1621 ACACTGAG 1680

QY 4807 ACAGCTTATAG 4866

Db 1681 ACAGCTTATAG 1740

QY 4867 AACATGATCCGTGTACAG 4926

Db 1741 AACATGATCCGTGTACAG 1800

QY 4927 CCTGACCTGTAG 4986

Db 1801 CCTGACCTGTAG 1860

QY 4987 CAG 5046

Db 1861 CAG 1920

QY 5047 GCCACTTAAG 5106

Db 1921 GCCACTTAAG 1980

QY 5107 GCACACCGT 5166

Db 1981 GCACACCGT 2040

QY 5167 CAGGCTTATTAAG 5226

Db 2041 CAGGCTTATTAAG 2100

QY 5227 CTGAGCTTGTAG 5286

Db 2101 CTGAGCTTGTAG 2160

QY 5287 AAG 5346

Db 2161 AAG 2220

QY 5347 TCCACGTCAAG 5406

Db 2221 TCCACGTCAAG 2280

QY 5407 GACATCCCCAGCTCAAG 5466

Db 2281 GACATCCCCAGCTCAAG 2340

QY 5467 GCCATCAG 5526

Db 2341 GCCATCAG 2400

QY 5527 GAGTTCAACACCATGAG 5586

Db 2401 GAGTTCAACACCATGAG 2460

QY 5587 GAGATCCTTGTGAG 5646

Db 2461 GAGATCCTTGTGAG 2520

QY 5647 CTGAG 5687

Db 2521 CTGAG 2561

RESULT 3

AY420221 5564 bp DNA linear GSS 12-DEC-2003

LOCUS Mus musculus HCM7151 gene, VIRUAL TRANSCRIPT, partial sequence,

DEFINITION genomic survey sequence.

ACCESSION AY420221

VERSION AY420221.1 GI:39776178

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 5564)

Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sainsky,J.J., Adams,M.D. and Cargill,M.

Interfing nonneutral evolution from human-chimp-mouse orthologous gene triles

JOURNAL Science 302 (5652), 1960-1963 (2003)

PUBMED 14671302

REFERENCE 2 (bases 1 to 5564)

Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sainsky,J.J., Adams,M.D. and Cargill,M.

Direct Submision

Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT This sequence as made by sequencing genomic exons and ordering them based on alignment.

FEATURES

source

1..5564

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

<1..>5564

/locus_tag="HCM7151"

ORIGIN

Query Match .37.1%; Score 2111.4; DB 29; Length 5564;

Best Local Similarity 58.5%; Pred. No. 0;

Matches 3140; Conservative 0; Mismatches 2155; Indels 77; Gaps 9;

QY 142 GCCAGAGGTTTCATTCACCTGT 201

Db 151 GACTGGGCTCATCTACCTGT 210

QY 202 GTCAATCGATTTPACAAAGCTCTCCAGGACCTTGAAAGTCTTGGTGAACGATGACAGAG 261
Db 211 GTGAACCGTATTTTACAAAGCTGTGCGGGAACTTAACCTGTCTGGGGCCCATGTGACAGGC 270
QY 262 CCGGACGAGACAAACCCAGTGTATACCAACCCCGCATGTGACAGACTTGAAAGAGCC 321
Db 271 CTTGTAAAGACATAGAAATGTCTACCCACCCCGCATGTAAAGTGTGCCCAATGTGT 330
QY 322 CTGACCAACCAACAAATGTCAACAGATCTCTCTATAGATCTTCAAGAGAAACAGGCTG 381
Db 331 CTGGGACACAGACAAAGTCAACAAAGCTTCTACTGTGATCTATGACCAACCCGCTG 390
QY 382 ATTGCTGTGGGACCTGTACCAAGGATCTGCAAGCTGTGAGGCTGAGAGACCTCTTC 441
Db 391 TTGGCTGTGTGCAATGCTCTACAGAGGCACTTGCAAGTCTTGAGTTGACATCTCTTC 450
QY 442 AAGCTGGGGAGACCTTATCATTAAGAGAGCACTATCTGTCAAGTGTCAAGAGAGGAG 501
Db 451 AAGCTGGGGAGACCGCACACCGCAAGAACACTACCTGTCAAGTGTCAAGAGAGGCTGAG 510
QY 502 TCACTCTTTGAGATGATC-----GTCTCTACAGCAACCTGGATGACAAAGCTGTCTAT 555
Db 511 AGCATGGCGGAGTGTCTCATTTGCTGGGCAACCGGCAAGGCAAGGCAAGCTCTTCGTG 570
QY 556 GCCACGGGAGTGTATGGGAAGCCGAGATTTTCCACCATCTTCAGCCGGAACCTGACC 615
Db 571 GGCAACCCATTCATGAGCAATGCAATGAAATATTTCCCACTGTCTCAAGTGTGCGGCTCAG 630
QY 616 AAGAACTGTAGGCGGATGCAATGTGCGGATGCTCTTCATATATGATGTGCGGCTG 675
Db 631 GCCAACAGGAAGATGTCTGATGATGTTGGCTTGTGTACAGAGATGATGTTGTATCAT 690
QY 676 ATGATTAAGATCCCTTTCGACACCTTCACCATCATCCCTGACTTGTATCTACTATGTC 735
Db 691 CAGCTTAAGATCCCTTTCAGACATTTGTCCAAAGTGTCCCGGCTTGTACATCTACTATG 750
QY 736 TATGTTTATAGCAATGAGCACTTGTCTACTTTTGAACCTTCAACCTGATGATGTGTCT 795
Db 751 TACAGTTTCGACGAGCGAGAGTGTGTACTACTCTCACTGACGCTGACAGCAAGCTT 810
QY 796 CCACAGAGCTCCACACCAAGAGAGAGAGTGTATCATCAAGCTCGTGAAGCTTTGCAAG 855
Db 811 ACCTGCGCCAGCGCGGCTGGGAGCACTTCTCACTCAAGATGTGCGGCTGTGCGT 870
QY 856 GAGGACACAGCTTCAACTCTATGTAGAGGTGCCATTGCTGTGAGCGCAAGTGGGCTG 915
Db 871 AATGACCCCAAGTCTCTACTCTTATGTGAGATTCCTCAATGGCTGTGAGAGGCTGCG 930
QY 916 GAGTACCGGCTCTGACAGGCTCTACTCTGCAAAAGCGGGGCGGCTGTGACAGGACC 975
Db 931 GAGTATCGGCTGTGACAGAGCGCTTACTGACCGCGGAGACAGGCGCTTGGCAGAGCAG 990
QY 976 CTTGAGTCCATCCAGATGATGACCTGCTCTTCAACGCTCTTCTCAAGGGGCGAGAGCG 1035
Db 991 CTGGGCTGTGAGTGTAGAGAGGTGTGTGTCAACGTTTCGACAGGGCGCAGAGAGAC 1050
QY 1036 AAAATTAATCCCTGATGATGATCGGCGCTGTGATCTTCACTTTGAAGCAGATTAATGAC 1095
Db 1051 CGCGTGAAGCCACCAAGAGATGCGCGCTGTGCTGTTCACACTGAGGAGCCATCAAGAG 1110
QY 1096 CGCATTAAGAGAGCGGCTGCAAGTCTTGTATCCGAGGAGAGGCAAGCTGACCTGAGCTG 1155
Db 1111 AAAATTAAGAGAGCGCATTCATCTGTCTACCGTGTGAGAGAGCAAGCTCTGCGCTGG 1170
QY 1156 CTCAAGGTGAAGACATCCCTGTGACAGAGTGGCTCTTAACCAATGACATTAATCTGT 1215
Db 1171 CTGTCTCAAGAGAGCTGGGCTGATCACTCGNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1230
QY 1216 GGCCTGACATGATGTCTCCCTGGAGGTGCGACATGATGTGCGTGAATTCCTGCTTC 1275
Db 1231 NNN 1290

QY 1276 ACGAGAGACAGGAGCCGATGACGTCGTATGCGCATATGTCFAAGAAACCACTCTG 1335
Db 1291 NNN 1350
QY 1336 GCTTTGTGGGACACCAAAAGTGAAGAGATCCGGGTGATGACCCAGAGGC 1395
Db 1351 NNN 1410
QY 1396 AAC-----GCCCTCAATATGACAGGTGAGGTGTGAGCCCGGCGAGTCTC 1446
Db 1411 AGCGGAGCGGCGGCTGGCTTATGAAGTGTGTGCTGACAGAGGGAACCTATCTG 1470
QY 1447 CGGATATGAGCTTCTCCAAAGAACAGAGCACTTACATCATGTGACAGAGGACCTC 1506
Db 1471 GGTACCTGTCTCTAGCCCAACCGCAAGTACTTATGCTATGACGAGAGACAGGTG 1530
QY 1507 ACCAGATCTCTGTGAGTCTGTGTGATCATATGAGCTGCGGAGTGTCTGCTCA 1566
Db 1531 ACGAGAGTCTGTGAGAAAGTGTGCAATATCTTGTGAGCTGTGTGCTGCTGCTG 1590
QY 1567 GAGGACCCCACTGTGCTGTGTGTGTGTCACAAACGTGCAACCGGAAGAGCGGTGT 1626
Db 1591 CGGAGCCCACTGTGCTGTGTGTGTCTGTGACAGCATCTGTCTAGGCAAGATGCTGT 1650
QY 1627 GAGCGGTCAAGAGACCCCGAGGTTGCTCGAGATGAAGCATGTGTCCGCTGACG 1686
Db 1651 GAGCGGAGAGAGAGCCCAAGCTTTGCTCTACCTGCTGACAGTGTGTGCACTTAACT 1710
QY 1687 GTCCATCCCAAAATATCTCCGTCTCAGTACAAACGCTGCTGTCTGAGACGTAC 1746
Db 1711 GTCCAGCGGCAATGT 1770
QY 1747 AATGTCCGAGAGTGTGAGTGTGCTGCACTGCACTTTGAGAGCTGTGAGATGAT 1806
Db 1771 AATGTCTGACCTGTGAGT 1830
QY 1807 GGGCTGT 1866
Db 1831 AGCATCTGGAAGATGCGCGATGATGCACTTCACTTGTGCGGGAAGTGTGCGGCC 1890
QY 1867 ATCATCAGAGAAATGAGGACCAACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1926
Db 1891 ATCAGAGGATCAGAGGAGACCAAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1950
QY 1927 GGCATGACTTGTGCGAGACCAAGCTTGTCTTCAATTAATGACAGTGTGTGTGTGTGTGT 1986
Db 1951 GGAAGAAATGT 2010
QY 1987 CTGTCTGCTGAGAGT 2046
Db 2011 CTGGAATGT 2070
QY 2047 CATGACCCCAAGACTGTCTCTTCAAGAAAGCGAGTGAAGCTGCGGAGAGCTGCGCC 2106
Db 2071 AACATGTGAGT 2130
QY 2107 CAGCTGTGCGAGT 2166
Db 2131 CAGATCTGTGCTTGT 2190
QY 2167 AAGGCAAGAACTTCCCAAGCCCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2226
Db 2191 GCTGTACGAGAACTGT 2250
QY 2227 ATTGAGGAGAGAGAGAGT 2286
Db 2251 ATCCAGGAGAGCGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2310
QY 2287 CAGAACACTCTTATCTCTATGAAGGATGAGATCAACAACTGCGCGGTGAGTGTGCA 2346
Db 2311 CAGAACTCTCGNN 2370
QY 2347 GTGTGTGGAATGGGCACTTCAACATGACCAACGCTCAGAAATAAGTTCACTCTAC 2406

[illegible][illegible]

[illegible]

RESULT 4	LOCUS	DEFINITION	ACCSSION	VERSION
AY420666	5520 bp	DNA	linear	GEN 12-DEC-2003
AY420666	Homc saliens HCM7300 gene, VIRUAL	TRANSCRIPT, partial sequence,		
AY420666	genomic survey sequence.			
AY420666.1	GI:39776623			

	KEYWORDS	GSS.
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euarchonta; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	Clark,A.G.; Glawinkel,S., Nielson,R., Thomas,P., Kejarival.A., Todd,M.A., Tanenbaum,D.M., Clevello,D.R., Lu,F., Murphy,B., Perreira,S., Wang,G., Zheng,X.H., White,T.J., Smnsky,J.J., Adams,M.D. and Cargill,M. Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios <i>Science</i> 302 (5652), 1960-1963 (2003) 14671302	
TITLE	(bases 1 to 5520)	
JOURNAL	Clark,A.G., Glawinkel,S., Nielson,R., Thomas,P., Kejarival.A., Todd,M.A., Tanenbaum,D.M., Clevello,D.R., Lu,F., Murphy.B., Perreira.S., Wang,G., Zheng,X.H., White,T.J., Smnsky,J.J., Adams,M.D. and Carrill,M. Direct Submission Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA This sequence as made by sequencing genomic exons and ordering them based on alignment.	
COMMENT	location/Qualifiers	
FEATURES	1..5520	
SOURCE	/organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" <1..>5520 /locus_tag="HCM7300"	
ORIGIN		
Query Match	35.4%; Score 2017.2; DB 29; Length 5520;	
Best Local Similarity	55.4%; Pred.No. 0;	
Matches 3018; Conservative	0; Mismatches 2393; Indels 39; Gaps 9	
DG	152 TCAATCACCTGGTGTAGTAGAGAAGACACAATTACTTTGGCGGCCTGAATCCGA 211	
OY	98 TTACCACCTGTCGTGTGCCAAGCAGTGAGTGTTCGTGGCGCATGAAACGAG 157	
OY	212 TTTCACAGCTCTCCAAGCAGCCTGAAGCTCTTGTTGTGACCCATGAAACAGGCCCCGACGAG 271	
DG	158 TCCTTAAGCTGGCCCCCAAAGCTGACTGACGTGCGGGCCCATAATGACCGGGCCCGTCGAG 217	
OY	272 ACAACCCCAAGTGTAACCAACCCCGCATGCTCCAGACTGCAATGAGCCCTGACCAACA 331	
DG	218 ACAAAGCTGCTGCTAACCAGCCCCCAGCAAGCGGTGTGTGCCACACGGCTGGSCCCCG 277	
OY	332 CCACATATGTCAACAAGATGTCTCTCATATGACTACACAGAGAAAACGGCTGATGCTGTG 391	
DG	278 TGACAAACATCAACAAGCTGCTGCTCATATGACTATGACGCGCCCGCCTGGTGGCTGGG 337	
OY	332 GAAGCTGTACCAAGAGGCACTGCAAGCTGCTGAGGCTGAGAGACTCTTTCAAGCTGGGCG 451	
DG	338 GGAGCATCTGGCAGGAGCACTGTCCAGCTTCTGTGCTGTGACGACCTCTTCAAGCTGGGCT 397	
OY	452 AGCCTTATCATAMGAAGAGCACTATCTGTGACGTGTCAACGAGAGCGGCTCAGTCTTTG 511	
DG	398 AGCGCACACACCGCAAGAGCACTACTGTGGGCGGCCAGAGGCCCGACTTCATAGAGCTG 457	
OY	512 GAGTATGTCTCTTCAACGCAACCTGTGATGACAACTGTCACTTCCACGCGAGTGGAT 571	
DG	458 GTGTATTGTGAGACAGGCGCCAGGGCCCAAGCACTGTGTTTTGGGCACTGCTGTGACG 517	
OY	572 GGAAGCCCGAGATATTTTCCCAACCATCTCCAGCCGAAAACTGACCAAGAACTCTGAGCGCG 631	
DG	518 GCAGATCGAAGATCTTCCCACTTGAGCTGCCGCAAGTCACTACATGATGTAAGACAGCG 577	
OY	632 ATGGATGTGCCGAGACGCTCTCCATGATGATGTGGCCCTCGATGATTAAGATCCCTT 691	
DG	578 CGGATGTGATGCTGTGCTGATACAGAGATTTTGTGCTCTCCAGATCAAGATCCCT 637	

QY 692 CGGACACCTTCCACCATCATCCCTGACTTGTATNTCTACTATGTCTATGTTTAAAGACTG 751
 DB 638 CAGACACGCTGCTCTTGTATCCCTGCTTGTGACATCTATCATCTACAGGCTTCTGACG 697
 QY 752 GCAACTTTTCTACTTTTGTGACCCCTCCAACTGAAATGCTGTCTCCACAGGCTCCACCA 811
 DB 698 CCTCTCTGCTGTACTTCTGACGCTGACGCTGAAACACCGACGACGCTGTTGACACAG 757
 QY 812 CCAGAGGACGAGGTGTATCATCCAAAGCTGTGAGGCTTTGCAAGAGGACACAGCTTCA 871
 DB 758 CGGAGCAAAATTTTTCAGCTCCAAAGATGTGTGATGTGCGGAGAGCTCAGAGTTCT 817
 QY 872 ACTCTATGTAGAGGTGCCAATGTGCTGTGAGCGCAGTGGGAGTAAACGCTGCTGC 931
 DB 818 ACTCATACGTGGAATTCCTCCATCGGCTGCTCTGCGCGCGCTGAGTACCGCTGTGTC 877
 QY 932 AGGCTGCTTACCTGTCCAAAGCGGAGCGGCTGTTGACAGACCTTGGAGTCCATCCAG 991
 DB 878 AGAGCGCTCCTGTGACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 937
 QY 992 ATGATGACCTGCTCTTCAACGCTTCTTCCAAAGGCGCAAGCGGAAATGAAATCCCTGC 1051
 DB 938 ATAGAGACGTCCTCTTCAACATCTTCTCTAGGCGCAGAAACCGGCGCACGCCACCC 997
 QY 1052 ATAGAGTCGCGCTGTGATCTTCAATCTTGAAGAGATTAATGACCGCATTAAGAGCGAG 1111
 DB 998 GGCAGACCATCTCTCTCTCTTCAACCTTCAACCAATCAATGCTCAATCCGCGCGCA 1057
 QY 1112 TGCAGTCTTGTATCCGCGCGAGAGGACGCTGACCTGCTGCTGCTCAAGTGAAGACA 1171
 DB 1058 TCAGTCTGCTGTATCGTGGGAGAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1117
 QY 1172 TCCCTGTGACAGTGGCTTTAAACATGACATTAATCTGTGCGCTGAGCATGAATG 1231
 DB 1118 TGCCTGTGACCAACCCCCCATGACATCAACGCACTTCTGTGGCTGTGTGTGAACC 1177
 QY 1232 CTCCTGTGAGTGTCCGACATGCTGCTGTGAATTCCTGCTTCAAGAGGACAGGAGCC 1291
 DB 1178 AGGCTCTGAGAGGCTGATGATGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1237
 QY 1292 GCAATGAGTGTGATGACATATGTCTACAGAACCACTCTGCGCTTTGTGGGACCA 1351
 DB 1238 GCATGCGCAGCTGCGCGCTTACACCTACCGCAGACCTGTGTGCTTCAATGTGACGC 1297
 QY 1352 AAATGCGCAGCTGGAAGAGATCCGGGTGATGGAACCCAGGAGCAACGCTTCAATATG 1411
 DB 1298 GCAAGCGAGCTTGAAGAG---NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1354
 QY 1412 AGACGCTGACGCTGTGAGACCCCGGCACTCTCGCGGATATGCTCTTCCAGAGACC 1471
 DB 1355 NNN 1414
 QY 1472 ACGAGCACTCTATCATGTGACAGAGGAGCTCACAGAGTCCCTGTGAGTCTGCTG 1531
 DB 1415 NNN 1474
 QY 1532 GTCAATATCAGAGCTGCGGAGAGTGTGCTGCTGACGCGACCCCACTGTGCTGTGTG 1591
 DB 1475 NNN 1534
 QY 1592 TGCCTGCAACACGCTGACCCCGAAGAGGCTGTGAGCGCTCAAGAGACCCCGAGAGT 1651
 DB 1535 TGTCTGCAACACAGGTGTGCTGCGCAAGGAGGCTGTGTGGGCTGTGCTGCTGCTGCTG 1594
 QY 1652 TTTGCTGAGAGTGAAGAGTGTGTGCGCTGACGCTCATGCCAATATCTCCGTCT 1711
 DB 1595 TTTGCTGAGAGTGAAGAGTGTGTGCGAGGTGCGGCTCCGACCAATGTGTCAAGTGA 1654
 QY 1712 CTCAGTACACGTGTGTGTGCTGCTGAGAGATGTCATGTCGCGGAGCTGTCAAGCTGGG 1771
 DB 1655 GGTCACTGTGGGTGTGAGTGTGACGCTGACCTGCAACGTGCAACGCTCAAGTGTGGGG 1714
 QY 1772 TCAACTGTACCTTTGAGACCTGTCAAGATGATGAG---CTGTGTGTGGGCAATCAGA 1828

DB 1715 TGAAGTGTGCTGTGAGAGCGGCGGAGAAAGAGGCGGTCTGTGCTGCTGCTGCTGCTG 1774
 QY 1829 TCCAGTGTACTTCCCTGACGCCAAGAGAGTGTGCGCGATCATCAAGAAATGGGAGCC 1888
 DB 1775 TGTCTGCGCCCTTCACTCCCTCCAGAGAGTCCAGACTCTTACAGAGGCGCATGTGGGCGCA 1834
 QY 1889 ACCATGTGTACAGCTTCAAGCTCAATCAAAAGAGACCGGCAATGACCTTGTGCGACAGCA 1948
 DB 1835 CCGCACTGTGGCTGTGAGCTTCTTCCAGAGAGACAGGCTGTGAGGCTTGTGCGGCTG 1894
 QY 1949 GCTTGTCTTCAATTTGAGCGGTCCACAAATTTGCTGTGCTGTGCTGTGCTGTGAGATCCAT 2008
 DB 1895 ACTTGTCTTCAATTTGAGCGGTCTCCAGAGTGTGCAATGTGTGTGTGTGCGAGCCCTT 1954
 QY 2009 ACCGCTGCTGCTGTGATTAATACCGGATGTGTGACACCATGACCCCAAGACCTGTGCT 2068
 DB 1955 ACCCTGTGCACTGTGTGATTAATGCTGCAACGCTGTACAGCGCCCTCCAGAGTGTCTCT 2014
 QY 2069 TCCAGAGGCGCAGTGAAGCTGCGAGAGTGTGCGCCCAAGCTGTGCTGCGAGTGAACAGA 2128
 DB 2015 TCCAGAGGCGCAGGCTGCAACAGCTGTGAGGCTGTGCGCTGAGATCTGTGCGAGGAGCC 2074
 QY 2129 TCTGTGTGCTGTGAGAGTGTATCAAGCTATCAAGCTGAAAGGCGCAAGAACTCCCGAGC 2188
 DB 2075 TCTGTATCCCGTGTGGGTGTATGACAGCTTCTTCACTGTGCGGCTTAAGAACTTACCTAGC 2134
 QY 2189 CCCAGTGTGAGCGCTGTGCTAAGAAATGATCTTCAACATTCAAGGCGAGCGAGCAGAG 2248
 DB 2135 CGAGTGTGCGCAGAAACATTAAGAGCTGTGTGCGGCTGTGAGAGGCGGCGCAGAGGAG 2194
 QY 2249 TGCCTGCTGTGCTGTGCTTCAACAGCTTCAAGCTGTGAGTGTGAGTGTGAGTGTGAGT 2308
 DB 2195 TGCCTGCTGTGCTGTGCTTCAACAGCAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2254
 QY 2309 AAGGATGTGAGATCAACAACTGCGGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2368
 DB 2255 AAGGTGTGTGAGATGTGACACGAGCTGTGCTTCTGCTGTGTGTGAGTGTGAGTGTG 2314
 QY 2369 ACATTTGACCAACCACTCAGAAATTAAGTTCAACCTTCAAGAGTGTGAGTGTGAGTGTG 2428
 DB 2315 CCATTAAGCAAGCTTCCAGCTTCCAGACCTTCCGTGCAAGTGTGAGTGTGAGTGTGAGTGTG 2374
 QY 2429 GCTGCGGCTGTGTCTTCAAGCTTCAACCACTTCAAGTGTGAGTGTGAGTGTGAGTGTG 2488
 DB 2375 GCTGTGCTGTGTCTTCAAGCTTCAACCTTCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2434
 QY 2489 GCAATGTGACCTGCTGCGCAGACCTGCTGCGCAGAGAGCAGTGTGAGTGTGAGTGTGAGTGTG 2548
 DB 2435 ACAAGTGTGACCTGCGACCCACCTGCGCGGCGCGGAGAGCAACATGTGATGACCTGAGCC 2494
 QY 2549 GTGCCAAAGCAATGTGCAAAACCCCGGATCAAGAGATTAATCCGAGTGAACAGGCGCC 2608
 DB 2495 AGAAGGACCCGGTGTGAGCAGCCCGGATCAAGAGATCAACCTTCTGTGAGGAGCTCA 2554
 QY 2609 GGAAGGGGAGCAACAGTCACTATCCAGAGGAGAACTGTGGCTGTGAAATTTGCGACA 2668
 DB 2555 AGAAGGAGGACCCGGTGTGACATGTGTGTGCAACCTTGTGAGTGTGAGTGTGAGTGTG 2614
 QY 2669 TGCCTTCCATGTCAAGTGTGCTGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2728
 DB 2615 TGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2668
 QY 2729 CTGCAAGACAGTGTGTGTGAGATGTGGGAGGCT---CAAGCCAGCTGAGTGTGAGTGTG 2785
 DB 2669 GTGTGTGAGAGATGTGTGTGAGATGTGAGATGTGAGATGTGAGATGTGAGATGTGAGAT 2728
 QY 2786 TGTGTGAGATGT 2845
 DB 2729 CCGTGTGAGCTGT 2788
 QY 2846 ATTATTTATGACATGACTCTCTCAATTTGAAGCCAGCTGTGGGCGCCATGTGCTGGAG 2905

Db 2789 ACAGCTTTGTGACCCCAACGTTTGACCAAGTGAATGCCAGCGGTGCGCGGTCCGGGG 2848
Qy 2906 GGAACCAAGTGAACATCAAGGACCAACTGAATGCCGGAAGCAAGTGTGTATGT 2965
Db 2849 GCACACGCTTACCATCTCAGGCGAGCTCTGTGATGTGGCAGAGGCTCACAGTACTG 2908
Qy 2966 TTGGAAGCAGCCCTGTCTCTTCCACAGCGAGATCTCATCTTCACTTGTGCAACCA 3025
Db 2909 TGAGGACAGCGAATGCGAGTTGTGAAGAGATGCGAAGGAGATGTGTGATCTCAC 2968
Qy 3026 CATCTCAGATGAGGTGCTAG--AGATGAAGTGTGTGCTGAGTGCAGAGGCCCAAGA 3082
Db 2969 CTCTCTCCACCTTGCGGCCCGCAGCGCCCATCACATTTGCCATTGACCGGGCTAACA 3028
Qy 3083 TTCA---CCAGACCTGGTCTTTCAATGTGAAGACCCACCATCGTGGATTGAGC 3139
Db 3029 TCTCCAGCCCCGGGCTCATCTACACCTACCTAGGACCCACCGTCAACCCGCTTGAGC 3088
Qy 3140 CAGATGAGGATTGTCAGTGGAACAACCCATCGCGTATGGGGACCCACCTGGACC 3199
Db 3089 CCACCTGGAGCATCATCAANN 3148
Qy 3200 TCATACAGAACCCCAAGTCCGTGCCAAGCATGGAGGAAAGAGACATCATATCTGTG 3259
Db 3149 NNN 3208
Qy 3260 AGGTTCTGAACGCTACTAGATGAAGCTGTAGCGGCCGCCCTCGCTCTGGGTCTGACC 3319
Db 3209 NNN 3268
Qy 3320 ACCAGTCAGACCTACCGAAGGCCCGAAGAGTTTGGCTTCATCTGGAACAAGTCCAGT 3379
Db 3369 NNN 3328
Qy 3380 CCCTGCTCATCTCAACAAGCAACTTCACTACTATCCCAACCCGGTGTGGAGCCT 3439
Db 3329 NNN 3388
Qy 3440 TTGTCCTCAGGAATCTGTGAGCTCAAGCTGCGACGCCCATCATCTTAAAGGCAAGA 3499
Db 3389 NNN 3448
Qy 3500 ACTGATCCGCTGTGGCTGGGGCAACGTGAAGTGAACCTACACTGTCTGTGGGG 3559
Db 3449 NNN 3508
Qy 3560 AGAAGCCGTGACCGGTGACCGTGTCAAGTGTCAAGTGTCTGTGAGTCCCCAACCTCA 3619
Db 3509 NNN 3568
Qy 3620 TCGGAGGCAACAAGTGAAGCGCGGTGCGGTGGCATGAGTACTCCCGGGAGTGTGT 3679
Db 3569 NNN 3628
Qy 3680 ACATGCCCCGAGACGCCGCTGAGCCCTGAGCCGCGCATGTCAAGCATGCAAGTGTGGCG 3739
Db 3626 ACATCTGGGAGAGGGGGGCTGACCTACCGGCGCATGATGGGGCTGCGGGCGGGGGTG 3685
Qy 3740 GCTTCTCATATTTTCAATGTGCGCGGTGCTCATTTGGCTATAAGCAAGTCCCGGAAA 3799
Db 3686 GAGTCTGTCTGTGCGCATCAAGCCGCTGTGTGGGTACAGCGGAAGCTTAGAGAG 3745
Qy 3800 GTGACCTCAGCGTGAAGCGCTGAGATGAGATGAGCAACTGGAAGTCCGTTGGGCC 3859
Db 3746 CGAACCCTACCTCAAGGTCTGACAGTGAAGTGAACAACCTGAAGTCCGTTGGGCC 3805
Qy 3860 TGAAGTGAAGAAAGCTTTTCCAGAGTGCAGAGAGATCAATGAAGTCAAGTGAAC 3919
Db 3806 TGAAGTGAAGAAAGCTTTTCAAGAGTGCAGAGAGATCAATGAAGTCAAGTGAAC 3865
Qy 3920 TGAATGAAGCGGAGTTCGTTCTGTGACTATGAACCTTACACATCGAGTGTGTCTTC 3979
Db 3866 TGAAGAGGTGCAAGTCCCTTCTGTGACTACCGGACTTACGCGGTGCGGTGTCTTCC 3925

Qy 3980 CAGGAATTGAAGAACCAACCTGTCTCCGGAGACTTGAAGTCCCGGGCTAACCGGAGAGC 4039
Db 3926 CGGGCATGAGGAGCCACCGGTGTCTCAAGAGCTG-----ATNNNNNNNNNN 3973
Qy 4040 GTGTGAAGAAAGGCTTCTGAGCTCTTCCGCGAGCTCATCAACAAGTGTCTCTGCT 4099
Db 3974 NNN 4033
Qy 4100 CCTCATCGGACGCTTGAATCCGAGCTAGCTTCTCATGCGGAGCCGTGGCAACGTGG 4159
Db 4034 NNN 4093
Qy 4160 CCTCATCATATGACCGTGTGAGAGCAAGCTGAGTACGCCACTGATGTGCTGAAGC 4219
Db 4094 NNN 4153
Qy 4220 AGCTCTGCGGACCTCATTTGACAAGACCTGAGAGCAAGAACCAACCTAAGCTGTGC 4279
Db 4154 NNN 4213
Qy 4280 TCAGAGAGACTGAGTCAAGTGTGAGAAATGCTGACCAATGGTTTACTTCTCTCT 4339
Db 4214 NNNNNNNNGACAGAGTCAATGTGCTGAGAAATGCTTACCACTGGTTCAGTTCCTGCTGC 4273
Qy 4340 ACAAGTCTCTCAAGAGTGTGTGGGGAGCCCTCTTCTCCCTGTCTGTGCAAGC 4399
Db 4274 ATAACTTCTGAAGAGTGTGTGGGGAGCTCTCTCTCTCTTACTGTGCTCATCAAGC 4333
Qy 4400 AGCAGTGAAGAGGGGCCCAATTGACGCTATCAAGGAGAGCGCGCTACTCTTGAAGC 4459
Db 4334 AGCAGTGAAGAGGGGCCCAATTGATGCCATCAAGGGAGAGGACGATACCTCCGAGCG 4393
Qy 4460 AGGACAAAGTCTATCCGACAGAGATGATGACTCAAAACCTGTGTCTGTGCTGTGAC 4519
Db 4394 AGGACAAAGTCTATCCGACAGAGATGATGACTCAAAACCTGTGTGCTGTGCTGTG 4453
Qy 4520 CAGCAATGCCAAGCGCCCGAGGCTCCAGTAAGATCTCTCACTGTGACACCATCTC 4579
Db 4454 CGGAGAAAGAGGAGCGGCCAGAGTCCAGTGAAGTGTCTCACTGTGACAGCATCAC 4513
Qy 4580 AGTCAAGAGAAATTTGTGATGCTCATCTTCAAGATGTGTCTTCCACCGGCCCA 4639
Db 4514 AGGCAAAAGATTAAGTGTGTGACACTGTGTCAAGGGCAATTCGTAACCTCAAGCTCCA 4573
Qy 4640 AAGCTGCAATTTGATCTGAGTGGGCGAGCAAGAAAGTGGGCAAGATGATCTTGCAG 4699
Db 4574 AAGCTGAGAACATGAGCTGAGTGGGCCCGCAGAGCGCGCATGATGATCTTCCAGG 4633
Qy 4700 ATGAAGACATCAACCAAGATTGAATGATTGGAACGACTGAACACTGTGCCACT 4759
Db 4634 ATGAAGATGTCAACCAAGATGAGTGTGATGAGAAAGGCTCAACTGTGGGCCACT 4693
Qy 4760 ACCAGTGCATATGTTCCGTGTGGGTGAGTATGTTCCAGAGGTGACAGCTTAACG 4815
Db 4694 ACCAGTGCATACAGTGTCTGTGTGGGTGAGTATGTTCCAGAAAGTGTCTGTCTTAACA 4753
Qy 4820 CAGTGAACATCCACGCTCTCAGAGCTCAAGAACTGCAAGTAAATATGAACAATGATCGGT 4879
Db 4754 TGGCCAACTCTTCACTTCAACCG--CTCCCTCAGCGGCTACAGAGCTGTGCCGA 4810
Qy 4880 ACAAGGAGCGCCCGACAGCTCTCCGCTCAAGCAACCTATGATCACTCTGACGAGGA 4939
Db 4811 CGGCAAGCAGCCCTATAGCTCTCGCTCAAGGAGCCCATGATTAAGCTTGAACAGAGA 4870
Qy 4940 GTGAAGTCAAGTGTGGCACTTATGAAGAACCAAGAGACGAGACCAAGAGAGGGGG 4999
Db 4871 CAGGACCAAAATTTGTGACCTGTGTGAAGAAACCAAGACATGCGCAACATGCGAGGGG 4930
Qy 5000 ACCGGGGAGCAAGATGTTGTGTAATCTTACCTGACCCGACTCTGTGCACTAAGGCA 5059
Db 4931 ACCGTGGAGCAAGATGTTGTGTAATCTTACCTGACCCGACTGTGCGCAACAAAGGGA 4990

QY	5060	AACAGCAAGAAATTGTGGATTAACCTCTTTGAGACCATTTGAGACGGCAACCGTGGCT	511.9
Db	4991	CACTCGAGAAATTGCTGGATGACCTCTTTGAGACAGTCTTACGACAGCCACCGGGGCT	505.0
QY	5120	CTGCGCTGGCCCTTGCCCATCAAGTACATGTTTGACTTCTCGATGAGCGAGGCTGTAAAC	517.9
Db	5051	CGGCGCTGGCCCTTGCCCATCAAGTACAGTTCCTCGATGAGCGAGGCGGACCGAC	511.0
QY	5180	ATGGCATTCATGACCCCGCACGCGCCATCCTGGAAAGAGCAATTGGCTGGCCCTGAGGT	523.9
Db	5111	GCCAGATCAGGAGACCCCGCATGTGGCCCAACCTGGAAAGAGCACTGGCTGGCGGTGGCT	517.0
QY	5240	TTTGGGTCACATGATCAAGAACCCGCAATTGTGTGTCATCCATTAAGAACAGCATCA	529.9
Db	5171	TCTGGGTCAAAATGTGATCAAGAACCCGCAATTGTGTGTCATCCATCAAGAACAGCATCA	523.0
QY	5300	CAGACGCGTGGCTCTCTGTGGTGGGTGAGACTTCATGAGACTCTTGCTCCACGTCAGAGC	535.9
Db	5211	CGGATGCTTGCTCTGTGGTGGGTGAGACTTCATGAGACTCTTGCTCTTCAATCCAGAGC	529.0
QY	5360	AACGGCTGGGCAAGGACTTGCCCTTCACAAAGCTGCTGTATGACCAGGACATCCCGAGCT	541.9
Db	5291	AACGGCTGGGCAAGGACTTGCCCTTCACAAAGCTGCTGTATGACCAGGACATCCCGAAT	535.0
QY	5420	ACAAGAATTGGGTGGAGAGATTACTCAGACATAGGGAAGATGGCCACCATCAGCGAGC	547.9
Db	5351	ACCAAGCTGGGTGGAGAGATTATTCAGACATTGGCAATGGCAATCCATCAGCGAGC	541.0
QY	5480	AAGACATGAACGATACCTGAGCTGAGCAGTCCCGGATGCACTGTAATGATTCAACACCA	553.9
Db	5411	AGGACATGAATGCTACCTGAGCTGAGCAGTCCCGGCTCAGCGCCGACATTCAGCGGCT	547.0
QY	5540	TGAGTGCACCTCTCAGAGATCTCTCTTAATGTTGGCAAAATACAGCAGAGAG	558.9
Db	5471	TGAGTGCCTCAACGAGCTGTATTTCTATGTACCAAGATACCGCAGGAG	552.0

```

RESULT 5
AY420668
LOCUS      AY420668
DEFINITION Mus musculus HCM300 gene, VIRTUAL TRANSCRIPT, partial sequence,
ACCESSION AY420668
VERSION    AY420668.1 GI:39776625
KEYWORDS   GSS.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 5447)
AUTHORS    Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
            Todd,M.A., Tamendamm,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferreira,S., Wang,G., Zheng,X.H., White,T.D., Smitsky,J.C.,
            Adams,M.D. and Cargill,M.
            Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
JOURNAL    Science 302 (5652), 1960-1963 (2003)
PUBMED     14671302
TITLE      2 (bases 1 to 5447)
AUTHORS    Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
            Todd,M.A., Tamendamm,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferreira,S., Wang,G., Zheng,X.H., White,T.D., Smitsky,J.C.,
            Adams,M.D. and Cargill,M.
            Direct Submission
JOURNAL    Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT    This sequence as made by sequencing genomic exons and ordering them
            based on alignment.
FEATURES   location/Qualifiers
            1..5447
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"

```

```

gene
<1..55447
/locus_tag="HCM7300"

Query Match      31.1%; Score 1767.2; DB 29; Length 5447;
Best Local Similarity 53.4%; Pred. No. 0;
Matches 2910; Conservative 0; Mismatches 2427; Indels 115; Gaps 9;

ORIGIN

150  TTTCATCATCCCTGCTGCTGATGAGAGAGACACATTTACTTGGGGGCGCTCAATCG 209
151  TCTGATCACTCTGAGCTGTGACCCGAGTACAGCTGGGGAGATGTTGTAGTCTCCGTAATCG 158
210  GATTTAAGAAGTCTCCAGCGACCTGGAAAGCTTGGTATACGATATGACAGGGCCGAGAGA 269
159  AGTTTCAAGTGGCTCCAAACCTTAACTTGAGCTACGGGCCCATGTCAAGGCCCATTTGA 218
270  GGACAACCCCAAGTGTATACCAACCCCGCATGTCTCAAGACCTGCATAGAGCCCTTGACC 329
219  GGACAATGCTCGCTGCTACCAACCTCTTAGATCCGTGTGTGCTCCACCGCCCTGTGCC 278
330  CACCAACATGTCAACAAGATGCTCCCTCATGTAGCTAACAAGAGAACAGGCTGATTTGCC 389
219  TGTGGAATATGAAACAAGCTGCTCTCTAAGATTAAGACCCCGTGGTTTGTAGCTTG 338
390  TGGAGCTCTGACCAAGGCATGTGCAGAGCTGTGAGGCTGAGGACCTTTCAAGCTGGG 449
339  TGGAGAGCATCTGGAGGGCATCTCCAGATTCCTGCTGGATACCTCTTCAAGTTGGG 398
450  GGAGCTTATATATAAGAGAGAGCACTATCTGTCAAGTGTCAACAGAGCGGCTCACTT 509
389  TGAAGCCCAATATGGCAAGAGCACTATCTGTCAAGGGCCCAAGAGGCTGATTCATGGC 458
510  TGAAGTATCGTCTCTCAAGCAACTGTGATGACAAAGTGTTCATTGCCACGGCAGTGA 569
459  TGGGCTCATTTGTTGAACAGGTCGAGGGGCTTAGCAAGTTGTTTGGGCAACGGTGTGA 518
570  TGGGAAGCCAGATATTTTCCACCAATCTCAGCGGAGAACTGACCAAGACTGTAGGC 629
519  TGGCAAGCTAGTACTTTCCACCTTGAAGTCCCGTAAGTCAATGCAATGAGAGAG 578
630  GGATGCAATGTCGGGATAGCTTTCATGATGATGATGATGATGATGATGATGATGATGAT 689
579  TGAATATGTTGAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 638
690  TTGGACACCTTCAACCATCATCCCTGATCTTGAATCATCATGATGATGATGATGATGAT 749
639  TTCAAGACATTTGCTTGAACCTGCTTGAATCATCATGATGATGATGATGATGATGATGAT 698
750  TGGCACTTGTCTACTTTTGAACCTTCAACCTGATGATGATGATGATGATGATGATGATGAT 809
699  TGCCCTCTCTGATGATCTTGAAGTGTGCACTGACACCAAGAGAGGCTGTGATAC 758
810  CACCAAGAGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 869
759  AGCAGGAGAAATTTCTTCAATCAACAGATGCTGTGATGATGATGATGATGATGATGATGAT 818
870  CAATCTTATGTAAGAGTGGCCATTTGGTGTGAGACGAGTGGGATGATGATGATGATGATGAT 929
819  CTACTCTTATGTAAGATTCCTTATTTGGTGTGCTCTGAGGTGATGATGATGATGATGATGAT 878
930  GCAAGCTCCTAATCTGTCACAAAGCGGGGCGGTGTTGGCAGAACCTTGGAGTGCATCC 989
879  GGAGAGTGCATCTGGCCAAAGCCAGGCGCTGTATGGCCAGGCTCTGGGTGTGCCAGC 938
990  AGATATATACCTGCTCTTCAACGCTTCTTCAAGGGCCAGAAAGGGAATGAATCCCT 1044
939  GCACAGAGATGCTCTTCAACATTTCTTCAAGGGCCAGAAAGACAGGCCAACCCAC 998
1050  GGATGATGAGGCTCTGATCTTCAATTTGAAGCAGATTAATGACCGCATTTAAGAGCG 1105
999  TGGGAGAGCAATCTTGTGCTTTTCAACCTGACAGCAATTAATGCTTCAATCCGCGCG 1055
1110  GGTGAGCTTGTATACGGGGGAGAGGAGACGCTGACCTGGCTGCTCAAGGTGAAGGA 1165

```

D	b		1059	AATCCAAATCGTGTCTAACGCTTGAGAGGGCAACGCTGACCCTACCTGGCTACTGAATTAAAGA	1118
Q	y		1170	CATCCCTGCAGCAGTCGCTCTTTAACCATTTGACGATPACTTCTGTGGCTTGACATGAA	1229
D	b		1119	GCTGCCCTGTGACTCAACACCCCTCTGCAGATCAATGAAAACTTCTGTGGGTATGTTGAA	1178
Q	y		1230	TGCTCCCCCTGGAGATGTCGCGCATGTGGTCGTGAATTCCGCTCTTCAACGGAGACAAGGA	1289
D	b		1179	TCAAACCATTTGGGTGGCCCTCACGTAATTTAGAGGGCTACCTTGTGGCTGACAGCACTGA	1238
Q	y		1290	CCGATGACGTCGTGTCATGCAATAATGTCATCAAGAACCATCTCTGGCCTTTGTGGCAC	1349
D	b		1239	TGGCATGGCCAGTGTGGCTGCCTACCTCACACAGCATCTGTGTGTTTTCATCGGTAC	1298
Q	y		1350	CAAAAAGTGGCAAGCTGAAGAAATCCGGGTGGATGAGCCGAGGGCAACGCGCTCCAGTA	1409
D	b		1299	ACGAGACGGTAAATTTGAAGAA-----	1319
Q	y		1410	TGAAGCGGTGAGGTGTGACCCCGGCCAAGTCTCTCGGGAATATGGCTTTCTCAAAGA	1469
D	b		1320	-----GNNNNNNNNNNNNNNNNNNNN	1339
Q	y		1470	CCAGAGCAACTACATCATGTCAAGAGGAGGTCACAGAGTCCTGTGAGTCTGT	1529
D	b		1340	NN	1399
Q	y		1530	TGTCAGTATCAGAGCTGCGCGCAGTGCCTTGCTCAGGCGACCCCACTGTGGCTGGT	1589
D	b		1400	NNNGTGTGGTG	1459
Q	y		1590	TGTGCTGCACAACAGTCGACCCGGAAGGACGGGTGTAGCGGTCCAAGAGGCCCGCAG	1649
D	b		1460	TGTGCTACAGCACAGTGTCTGCCCTGAAGGGGCTGTTCAGTGTGCTCTGCGCCACATGG	1519
Q	y		1650	GTTTGCTCGAGATGAAGCAGTGTGTCCGCTGACGTCATCCCAATATCTCCGT	1709
D	b		1520	CTTTGCAGAAAGCTGAGCAATGATATACAGTGGGGTCCGTCCCAATATGTGTAGT	1579
Q	y		1710	CTTCAGTACAACGTCGTGTGTGTCCTGAGACGATCAATCCCAGGCTGTAGCTGG	1769
D	b		1580	GACATCTCTGGGGGTGCAGTCGACCGTGGCATGGCAACGTGCAAGACCTCAATGTGGG	1639
Q	y		1770	CGTCACTGCACCTTTGAGACCTGTCAAGATGGATGGCTGCTGTGTGGCAAT---CA	1826
D	b		1640	TGTGAGCTGTCTCTTTGAGAGGTGACGAAAGTGAAGGCTATTTGCTGCGCCCTGTGAGA	1699
Q	y		1827	GATCAGTGTACTCCCTGCAAGCCAAAGAGTGCCTCCGATCATCAGAGAAATGGGA	1886
D	b		1700	GTTAGATGTCTTCAACATCCCTTCAAGAGCTCCAGACACTTACAGAGGGCATGGGCG	1759
Q	y		1887	CAACCATGTCTGACAGCTTCAGCTCAAAATCAAAGGAGCCGCGATGACTTTCGCCAGAC	1946
D	b		1760	CACATCACTGTGGGCTTCAGCTGTCTCATGAGACGTGATGAAGTGTGCTGGGCT	1819
Q	y		1947	CAGCTTGTCTTTCAAAATTCAGCGGTCCAAATTCGTGCTGTCCTGCTGAGAGTCC	2006
D	b		1820	TGACTTGTCTTTCAACAATGAGTGGCCCTCAAGTCGTATATGTCGTGTGTGGAGGCC	1879
Q	y		2007	ATACCGCTGCACATGCTGTAAATACCCGGATGTCTGCACCCATGACCCCAAGACCTGTC	2066
D	b		1880	TTACCCCTGCACATGCTGTAAATACCGTCAATGTGTGACAGACCCACAGATGTCT	1939
Q	y		2067	CTTCAGGAAGGCGCAGTGAAGTCCGCCAGAGACTGCCCAAGCTGTCTGAGTGAACA	2126
D	b		1940	TTTCAGGAGGGCAGGGTCCACAGCCTTAGAGGCTGCCCTGAAGTCCGCTCAAGGGGA	1999
Q	y		2127	GATCTGTGTCCTGTGAGTGTATCAACGCTATACGCTGAAGGCCAAGAACTCCCCCA	2186
D	b		2000	CCTCTTAATCTGTGGGTGTATCAAGCTCTTAATCTTGGGGCTTAAGAACTTGCCACA	2059
Q	y		2187	GCCCCAGTCTGGGACGCTGCTACGAATCTCTCAACATTCGAGGCAACGAGACGG	2246

[illegible]

```
QY 3318 CCACAGTCAGACCTGACCGAGAGGCCGAGAGATTGGCTTCATCTCGACAACGCTCCA 3377
Db 3194 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3253
QY 3378 GTCCCTGCTCATCTCAACAAGCACTTCACTACTATCCCAACCCGGGTGTTGAGGC 3437
Db 3254 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3313
QY 3438 CTTTGGTCCCTCAGGAATCCTGAGCTCAAGCTGGACGCCCATCATCTTAAAGGCGCA 3497
Db 3314 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3373
QY 3498 GAACTGATCCGCCCTGTGGCTGGGGGCAAGTGAAGTGAACCTACCTGTGTGGTGG 3557
Db 3374 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3433
QY 3558 GGAGAACCGTGACCGGTGACCGTGTCAAGTGTCAAGCTGCTGTGAGTCCCCCACT 3617
Db 3434 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3493
QY 3618 CATCGCAGGCAACAAGTATGGCCCGTGTGGTGGCATGAGATCTCCCGGGGATGCT 3677
Db 3494 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3556
QY 3678 GTACATTTGCCCGGACAGCCCGCTCAGCTGCCCGCATCTCAGCATCGCAGTGGCTGG 3737
Db 3551 GCATCATCTGCTGATCGGGCACTGACCTTACAGCTATGTGGGGCTAGAGACAGGGGG 3610
QY 3738 CGGCTTCCTCATTTTCACTGTGGCCGTCTCATTTGCTTAAACGCAAGTCCCGGCA 3797
Db 3611 CGGGCTCTTGCTGCTGGCATCATGTTGTGTGGCTTCAAGCGTAAACATCAAGGA 3670
QY 3798 AAGTGACCTCAGCCTGAGGGGTGAGATGAGATGAGCAACTGAGTCCCGTGGGC 3857
Db 3671 TGGAGCCGCACTTAAGCGCTTACGCTAACATGAGCAACTGGAGTCTCGTGTGGC 3730
QY 3858 CTTGAGTGAAGAAACCTTTGCGAGCTGCAAGCGACATCATGAGCTGACCAAGTGA 3917
Db 3731 CTTGAGTGAAGAAAGCTTTGCTGAGCTGCAAGCTGATATCAATGAATCAACAACA 3790
QY 3918 CCGAGTGAAGCGGATTCCTGCTGAGCTATGAATCAATGACATGCGAGTCTGCT 3977
Db 3791 CATGATGTGTCAATCCCTCTTGTGATCGGACCTAATGCTGTGGGTCTTT 3850
QY 3978 CCCAGAAATTGAAGCAACCTGTCTCCGGGACCTTGAAGTCCGGGCTACCGGAGGA 4037
Db 3851 CCCTGGCATGAGGCTCAACCAGTGTCTCAAGGAGCTG-----ATTNNNNNNN 3898
QY 4038 GCGTGTGAGAAAGGCTGAAGCTTTCGCCAGCTCATCAACAAGGTTCTCTGCT 4097
Db 3899 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3958
QY 4098 GTCTTCATCCGACGCTTGAATCCAGCGTACCTTCTCATGCGGACCGTGGCAAGCT 4157
Db 3959 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4018
QY 4158 GGGCTACTCATATGACCGTGTGCGAGCAAGCTGAGTACGCCACTGATGTGTGA 4217
Db 4019 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4078
QY 4218 GCAGCTGCTGCGACCTCATTTGACAAGAACTGAGAGCAAGAACCACTTAAGCTGCT 4277
Db 4079 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4138
QY 4278 GCTCAGAGAGACTGAGTCAAGTGGTGAAGAAATGCTGACCAATTGGTTTATCTTCTCT 4337
Db 4139 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4198
QY 4338 CTACAAATTCCTCAAGAGTGTGTGGGAGGCCCTTCTCTCTGTTCTGTGGCATCA 4397
Db 4199 GCATTAATTCCTGAAGAGTGGCGCGGGAGGCCACTTCTCTGTGACTGGGCATCA 4258

QY 4398 GCAGCATGTGAGAAAGGCCCATTTGAGCCATCAAGGGGAGAGGCCCGCTACTCTTGG 4457
Db 4259 GCAGCATGTGAGAAAGGCTCTATTGATGCATTAAGGCGAGAGGCCGCTTACTTTAAG 4318
QY 4458 CGAGCAAGCTCATCCGCGAGAGATTGACTAACAACCTGTGCTGAGCTGTGTCAG 4517
Db 4319 TGAGCAAGCTCATCCGCGAGAGATGACTATAAGACCTGACTCTGATTTGTGTGG 4378
QY 4518 CCCAGCAATGCCAAGAGCCCGAGGTCCTCAAGAAATCCCACTGTGACACATGAC 4577
Db 4379 CCGGAGAGGAGGAGGATGCCCGAGTCCCTGTGAAGGTTCTTAAGTGTGAATGATCAC 4438
QY 4578 TCAGTCAAGAGAAATTCGTGATCCATCTTCAAGAAATGTGCTGTGCCACCGGCG 4637
Db 4439 CAGGCCAAGAACAAGTGTGATACTGTGTACAAGGTATTCATCTCAAGCGCC 4498
QY 4638 CAAAGCTGAGATATGATCTGAGTGGCGACAAGAAATGGGGCAAGATATCTTGA 4697
Db 4499 CAAAGCTGAGATATGATTTGAATGGCGCGAGGCCCATGGCCGATCATCTTCCA 4558
QY 4698 GATGAAGACATCAACCAAGATTGAGATGATGAAAGCACTGAACAACCTGGCCA 4757
Db 4559 AGATGAGACATCACTACAAGATGATGTGATGAGAGAGGTCATCTGATGGCCA 4618
QY 4758 CTACAGGTGCCAGATGTTCCGTGTGGCATTAATGTCTCAAGCAAGTGAACCTTAA 4817
Db 4619 CTACAGGTACTGATGTGTTCTTATGATGACGTGGCCCAACAAGTGTGCTATTA 4678
QY 4818 CGAGTGAACAATCCACCGTCTCAGAGACCTCAGCAAGTAAATATGAAACATGATCCG 4877
Db 4679 CARGCCATCTGTTACCTTACCTTACCGG---TCACTTATGTGCTGATGAGAGCTGTCCG 4735
QY 4878 GTACACGGGAGGCCCGGACAGCTTCGCTCAAGCAACCTATGATCACTCTGACCTGA 4937
Db 4736 TGCTGCAGACAGCCGAGTATGCTTCGTTCCAGAGCACTATGCTACCGCTTACCAAGA 4795
QY 4938 GAGTGAAGTGAAGTGGACCTGATGAAAGAACCAAGCAAGAGCAAGAGAGAGG 4997
Db 4796 GGCAGGTACCAAGCTGTGGCACTGTGTGAGAAACCAAGCACATGATCAACGAAAG 4855
QY 4998 GGAACGGGGAGCAAGATGTGTCTGAATCTAAGTCCGACCTCTGGCCACTAAGG 5057
Db 4856 AGACCGCGGAGCAAGATGTGTCTCAAGAAATATCTCAAGGCTCTGGCCCAAGAGG 4915
QY 5058 CACATGCAAGAATGTTGTGATGACCTTTGAGACATTTGAGACGCGCACACGCTGG 5117
Db 4916 CAGCTTCAGAAAGTTTGTATGATGATCTTTTGAATCTGTGTGATGATGACCAAGCGGG 4975
QY 5118 CTCTGCCCTGCCCCGTGCAATCAAGTACATGTTTGAATCTTCTGATGAGCAGCTATA 5177
Db 4976 CTACGCTTACCTCTGGCCATCAAGTACATGTTTGAATCTTCTGATGAAAGAGCTACCA 5035
QY 5178 ACATGGAATTCAGAACCCGACGTCGSCATTAAGTGAAGAAATGCTGCGCCGTGAG 5237
Db 5036 GCGCAGATGATGACCTTATGTGTGCTCACTGTGAAGAGCACTGTTACTCTGCG 5095
QY 5238 GTTTGGTCAACATGATCAAGAAACCCGAGTTTGTGTTGATCATTCATTAAGACAT 5297
Db 5096 TTTCTGGTGAATGTATCAAGAAATCCGCAATTTGTGTTGATCATTCATTAAGATAGCAT 5155
QY 5298 CACAGAGGCTGCTCTCTGTGTGTGCTCAGACCTTCAATGATGATCTGCTCCAGCTGAG 5357
Db 5156 CACAGAGGCTGTTTCAAGTGTGTGCTCAGACCTTATGATGATCTGCTGTTTCAATGAG 5215
QY 5358 GCAACGCTGGGCAAGACTCGCCCTCAACAAGCTGCTATGCAAGACATCCCGAG 5417
Db 5216 GCAACGCTGGGCAAGACTCACTTCAACAAGCTCTCTATGCAAGATATTCCTCA 5275
QY 5418 CTACAAATTTGGTGAAGATTTTACTAGACATAGGAAATGTCAGCAATCAAGCA 5477
Db 5276 TTACAAAGCTGGTGAAGATTTACTATGAGATATGCAAAAGTGAATCATCATGATGA 5335
QY 5478 CCAAGACATGAACCATACTGCTGAGCAGTCCCGGATGACATGAATGATTCACAC 5537
```


DB 5336 CCAGACATGACGCGCTACTTAGTGAGACAGTCCGCGCTCATGCTATGACTTCAATGT 5395
QY 5538 CATAGTGCATCTTCAGAGATCTTCTCTATGTGGGCAATATACGAGAG 5589
DB 5396 CTAATGTCACCTACGCGAGCTCTACTTCTATGTACCAAGTACCGTCAAGAG 5447

RESULT 6
AY420667 5520 bp DNA linear GSS 12-DEC-2003
LOCUS AY420667
DEFINITION Pan troglodytes HCM7300 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY420667
VERSION AY420667.1 GI:39776624
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 5520)
Clark,A.G., Gnanowsky,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civeello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene titers
Journal
Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
AUTHORS 2 (bases 1 to 5520)
Clark,A.G., Gnanowsky,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civeello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence as made by sequencing genomic exons and ordering them
based on alignment.
FEATURES
source Location/Qualifiers
1..5520
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>5520
/locus_tag="HCM7300"

ORIGIN
gene
Query Match 29.5%; Score 1661.4; DB 29; Length 5520;
Best Local Similarity 46.8%; Pied. No. 0;
Matches 2552; Conservative 0; Mismatches 2859; Indels 39; Gaps 9;

DB 152 TCATACCTGCTGCTGATGAGAGACAGACATTTACTGGGGCCCTCAATCGA 211
QY 98 TTACCACTGCTGCTGACACCGGATGACCTGGGAGGTGTNNNGGNNNNNNNNNGAG 157
DB 212 TTACCACTGCTGCTGACACCGGATGACCTGGGAGGTGTNNNGGNNNNNNNNNGAG 271
QY 158 TCTTTAAGTGGCCCCCACTGACTGAGCTGGGGCCCATGTCAAGGGCCCGTNNNG 217
DB 272 ACAACCCAGAGTTTCCACCGCGATGCTCCAGACTGAGTGGCCCTGACACCA 331
QY 218 ACAACGTCGCTCTACCCGCCCAAGCATGCGCTGTGTGCCACCGCTGGCCCG 277
DB 332 CCAACATGTCAACAGATGCTCTCATAGACTCAAGAGAAACAGGCTGATGCTGTG 391
QY 278 TGACAAACATCAACAGATGCTCTCATAGACTCAAGAGAAACAGGCTGATGCTGTG 337
DB 392 GAGACCTGTACCAAGATGCTGCAAGCTGTGAGGCTGAGAGACCTTTCAAGCTGGGG 451
QY 338 GCAAGCATCTGACAGGGGATGCGAGTCTGCGCTGTGAGCAAGCTTTCAAGCTGGGG 397
DB 452 AGCTTATCATTAAGAGAGACATATCTGACAGTGTCAACAGAGCGGCTAGTCTTTG 511

DB 398 AGCGACACACCGCAAGAGGCAANNCTGTCCGGGGGCCAGAGGCCCATCTCATGGCTG 457
QY 512 GAGTATGCTCTTCCATACAGCAACCTGGATGACAGCTGTTCATTTGCCACGCAATGATG 571
DB 458 GTNNNNNNNGTGAAGAGAGGCCAGAGGCCAGCAAGCTGTTTGGGCACTGCTGTNNNNN 517
QY 572 GGAAGCCCGAGTATTTTCCACCATCTCCAGCGGAAACTGACCAAGACTGTGAGCGG 631
DB 518 NNN 577
QY 632 ATGCGATGTTGCGGATGCTCTTCATGATGAGTGTGCTGCTGATGATTAAGCTCTT 691
DB 578 NNN 637
QY 692 CGACACCTTACCATCATCTGCTGATGATATCTACTATGCTATGCTTTAGAGTG 751
DB 638 CAGACAGCTGCTCTTGTACCCCTGCTTTGACATCTACTACATCAAGCTTCTGTCAGCG 697
QY 752 GCACTTGTCTACTTTTGAACCTTCAACCTGATGATGATGATGATGATGATGATGATG 811
DB 812 CCAAGAGAGAGTGTATATATCAATCCAGCTGTGAGGCTTTGCAAGAGAGACAGCTTCA 871
QY 758 CGGCGAGAAATTTTCAAGTCCAGATGATGATGATGATGATGATGATGATGATGATG 817
DB 872 ACTCCATATGAGAGGAGGCCATGAGCTGAGAGCGAGGAGGAGTACCGCTGCTG 931
QY 818 ACTCATACGTGATATTCCTCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 877
DB 932 AGCTGCTTACCTGTCCAAAGCGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 991
QY 878 AGAGCGCCACTGCGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 937
DB 992 ATGATGACCTGCTTCTCAACCTGCTTCCAAAGGCGGAGGAGGAGGAGGAGGAGGAG 1051
QY 938 ATGAGGACATCTTCTTCAAGATCTTCTTCAAGGCGGAGGAGGAGGAGGAGGAGGAG 997
DB 1052 ATGAGTGGCCCTGTGATCTTCAATGAGATTAATGACCGCATTAAGAGCGGC 1111
QY 998 GGAAGACATCTCTGCTTCTTCAACCTGCTTCAACCTGCTTCAACCTGCTTCAACCTG 1057
DB 1112 TGCACTTGTATACCGGGGAGAGGAGAGGAGAGGAGAGGAGAGGAGAGGAGAGGAGAG 1171
QY 1058 TCCAGTCTGCTATCGGGGAGAGGAGAGGAGAGGAGAGGAGAGGAGAGGAGAGGAGAG 1117
DB 1172 TCCCTGACAGAGTGCCTCTTAAACATGAGATTAATCTGTGCTGAGATGATG 1231
QY 1118 TGCCCTGATCAACACCCCATGAGATCAACGCAACTTCTGTGGGCTGTGTTGAAAC 1177
DB 1232 CTCCTGAGAGTGTCCGATGATGAGTGTGAGATTTCCCGTCTTCAAGGAGAGAGGAGAG 1291
QY 1178 AGCTCTGAGAGGCTGATGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1237
DB 1292 GCATGACGTCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1351
QY 1238 GATGAGCAGGTGTGCGGCTTACACTACCGCAGAGACTGTGNNNTTTCATGTGGCAGCG 1297
DB 1352 AAATGCGAAGCTGAAGAAGATCCGGGTGATGAGCCAGAGGAGAGGAGGAGGAGGAG 1411
QY 1298 GCAAGCGCAGNN 1357
DB 1412 AGACGTGAGGTGTGAGCCCGGCGCAGTCTCTCGGAGATATGCTTCCAAAGAGC 1471
QY 1358 NNN 1417
DB 1472 ACGAGCAACTTATCATGTATGATGAGAGGAGAGCTCACAGAGTCCCTGTGAGTGTG 1551
QY 1418 NNN 1477
DB 1532 GTACGTATCAGAGTGGCGGAGTGTGCTGTGCTGAGGAGAGCCCATCTGTGCTGTG 1591
QY

QY	3740	GCCTCCCTCATCTTTTCATCGTGGCCGTGTGCATCTGGCTCTTAAACGCAATGCCGCGAA	3799
Db	3886	GGCTCTGCTGCTGCGCATCACGGCCGTGCTGGTGGCTTCAACGCGAACAACNNNNNACG	3745
QY	3800	GTCACCTCACGCTGAAGCGCGCTGCGAGATGCAGATGCAACACTGAGAGTCCCTGTGGCC	3859
Db	3746	CGGACCTTACCCNNNAACGCTCTGCAGCTNNNNATGCAACAACCTGGAGT-----	3793
QY	3860	TGAGTGCAGGAAGCTTTGCCGAGCTGCAGACGACATCTAGACTGACCAAGTAC	3919
Db	3794	NN	3853
QY	3920	TGATGAGCCCGGATTCCGTTCTCTGGACTATAGACTTACACCATGCGGCTGCTCC	3979
Db	3854	NN	3913
QY	3980	CAGGAATTGAAGCAACCTGTCTCTCGGACCTTGAGTCCCGGCTACCGGACGAGC	4039
Db	3914	NN	3973
QY	4040	GTTGAGAAAAGGCGCTGAGCTCTTCCGCCACTCATCAACACAAGTGTCTGCTGT	4099
Db	3974	NN	4033
QY	4100	CCTTCATCCGACGCTTGAGTCCGACGTAGCTTCTCATGCGCGACCGTGACCACTGG	4159
Db	4034	NN	4093
QY	4160	CCTCATCTCATGACGCTGCTGCAGAGCAAGCTGAGTACGCACTGATGTGCTGAAGC	4219
Db	4094	NN	4153
QY	4220	AGCTGCTGCGCAGCTCATGTACAAGAACTGAGAGAGAAACACCTTAAGCTGTGC	4279
Db	4154	NN	4213
QY	4280	TCAGAGAGACTGAGTCACTGCTGCTGAGAAAGATGCTGACCAATTGGTTACTTCTCTCT	4339
Db	4214	NNNNNNNACACAGAGTCACTGGCTGAGAAAGTCTTACCACTGTGTACCGTTCGTGTGC	4273
QY	4340	ACAAGTTCCTCAAGAGTGTCTGTGGGAGCCCTCTTCTCCCTGTTCTGTGCACTCAAGC	4399
Db	4274	ATAAGTTCCTAAGAGAGTGTCTGTGGGAGCCCTCTTCTCTTACTGTGCACTCAAGC	4333
QY	4400	AGCAGATGGAAGAGGGCCCCCATTTGACGCATCACGGGGAGGCCCGCTACTCTTGAGCG	4459
Db	4334	AGCAATGGAAGAGGGCCCCCATTTGATGCTATCACGGGGAGAGCAAGATCTCTTGAGCG	4393
QY	4460	AGGACAACTCATCCGACAGACAGATTTGACTCAAAACCTGCTCTGAGCTGTGTACGC	4519
Db	4394	AGGACAACTCATCCGTGACAGAGATCGACTCAACAGCACTGACCTTCACTGCGGTGTCTC	4453
QY	4520	CAGACAAATGCCAACAGCCCCCGAGGTCCCAAGTAAGATCTCTCACTGTGACACATCACTC	4579
Db	4454	CGGAAGAACGAGGGCAGCCGCCAGGTGCCAGGTGAAGGTTCTCACTGTGACACATCACTC	4513
QY	4580	AGGTCAAGAGAGAAGATTGTGATGCCATCTTCAAGAAATGTGCTTGTCTCCACCGGCCCA	4639
Db	4514	AGGCCAAAGATTAGCTGTGTGCACTGTGTACAAAGGGAATTCCTGACTCCAGCGTCCCA	4573
QY	4640	AAGCTGCAGATTTGATCTGAGATGGCGCAACAAGAAATGGGGCAAGATGATCTTGACG	4699
Db	4574	AAGCTGAGGACATGAGCTCTGGAGTGGCGCAAGGGCCGATGACTGTGCATCACTCTCAAGG	4633
QY	4700	ATGAAGACATCCACCAAGATTGAGATGATTGGAACGACTGAACAACACTGGCCCACT	4759
Db	4634	ATGAAGATGTCAACCAAGATCGAGTGTGACTGGAAGAGGCTCAACTCACTGGGCCACT	4693
QY	4760	ACCAAGTGGCAATGTTCCGATGTGGCAATTGATGCCAAGAGAGTGCAGGCTATTAAG	4819
Db	4694	ACCAAGTGCAGACGGTTCCTTGGTGGCAATTGTGTGCCCAACAAATGTCTGCTATTAAC	4753
QY	4820	CAGTGAACAATCCACCGTCTCCAGACCTCAGCAAGTAAATATGAACAATGATCCGCT	4879

Db		4754	TGGGCAACTCTTTCACCTTACCAGCG--CTCCCTCAGCCGCTACGNNNNNNNNNNNNNNN	4810
Oy		4880	ACAACGGGCACGCCCGACAGCGCTCCGCTCACGGACACACTATGATCATCTCCTGACGTGAGA	4939
Db		4811	NN	4870
Oy		4940	GTCGAGTCAAAGATGTGCGACCTTAGTAAGAACCAACGACACGGAGACCGAAGAGAGGGGG	4999
Db		4871	NNNGCACCNNAATTGTGGCACCTGGTGNNNNNCCAACGACCATGCGACATGCGACGGGGG	4930
Oy		5000	ACCGGGGGAGACAGATGGTGTCTGAATCTACCTGACCCGACTCTGGCCACTAAGGGCA	5059
Db		4931	ACCGTGGCAGNNNNNTGGTCTCNNAGATCTACCTGNNNNNNNNNNNNNNNNNNNNNNNGCA	4990
Oy		5060	CACCTGCAAAATTTGTGATGACCTCTTTGAGACATCTTCAGACGGCACACCGTGGCT	5119
Db		4991	CACCTGCAAAATTTGTGATGACCTCTTTGAGACAGTTCAGACAGCCACCGGGGCT	5050
Oy		5120	CTGCCCTGCCCCCTGGCCATCAAGTACATGTTTGACTTCTCGATGACAGGCTGATTAAC	5179
Db		5051	CGGCGCTCGCCCTGGCCATCAAGTACATGTTTGCATCTCTGGATGACAGGCGGACCGAG	5110
Oy		5180	ATGCGATTCAATGACCCGCAAGTCCGCAATACCTGGAAAGACAAATTCCTGCTCCCTGAGGT	5239
Db		5111	GCCGAGATCAGCGACCCCGAGCTGGCGCACACCTGGAAAGACAAATTCCTGCTCCCTGAGGT	5170
Oy		5240	TTTTGGTCAACATGATCAAGAACCCGAGATTTGTGTCATCTCATPAGAACAGCATCA	5299
Db		5171	TCTGGTGAATGTGATCAAGAACCCAGATTCGATTCGATCTCAACAGAACAGCATCA	5230
Oy		5300	CAGACGCTTGCTCTCTGTGTGTGCTCAAGCTTCATGAGACTCTTCTCTCCACGTCAAGC	5359
Db		5231	CGGATGCTGCTGT	5290
Oy		5360	ACCGGCTGGGCAAGGACTCGCCCTCCAAACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGT	5419
Db		5291	ACCGCTGGGCAAGGACTCGCCCTCCAAACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGT	5350
Oy		5420	ACAAAGATTGGGTGGAAGGTATTACTACGACATAGGGAAGATGCGACCTCAGCGACG	5479
Db		5351	ACAAAGCTGGGTGGAAGGTATTACTACGACATAGGGAAGATGCGACCTCAGCGACG	5410
Oy		5480	AAGACATGAACGCACTACCTGGCTAGAGAGTCCCGATGACATGACATGATGATTCAACACCA	5539
Db		5411	AGGACATGAACGCACTACCTGGCTAGAGAGTCCCGATGACATGACATGATGATGATTCAACACCA	5470
Oy		5540	TGAGTGCATCTCTCAGAGATCTTCTTCCTATGTGGGCAATTAAGAGAGAG	5589
Db		5471	TGAGTGCATCTCTCAGAGATCTTCTTCCTATGTGGGCAATTAAGAGAGAG	5520
RESULT 7				
LOCUS	AY420220	5341 bp	DNA	linear GSS 12-DEC-2002
DEFINITION	Pan troglodytes HCM1751 gene, VIRUTUAL TRANSCRIPT, partial sequence.			
ACCESSION	AY420220			
VERSION	AY420220.1	GI:39776177		
KEYWORDS	GSS.			
SOURCE	Pan troglodytes (chimpanzee)			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.			
REFERENCE	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,			
AUTHORS	Todd,M.A., Tenebaum,D.M., Clevello,D.R., Lu,F., Murphy,B.,			
	Ferriera,S., Wang,G., Zheng,X.H., White,T.D., Shinsky,J.J.,			
	Adams,M.D. and Cargill,M.			
TITLE	Infering nonneutral evolution from human-chimp-mouse orthologous			
	gene trics			
JOURNAL	Science 302 (5652), 1960-1963 (2003)			
PubMed	14671302			

QY 2196 TGGGCAAGCTGCTACGATGATCTCTCAATTCAGGGCAGGAGCAGAGTCCGCG 2255
DB 1861 AGGCAAGCTGATATGAGTGCCTCTCCACATCCCGGGAGCGCCGCGCTGTCAACCC 1920
QY 2256 CCGCGCTTCAACAGCTCCGAGGTACAGTGCAGAAACACTC--TTATTCATGAAG 2312
DB 1921 CTTGCGCTTCAACACTCCAGCTCCAGTGCAGAAATTCCTGANNNNNNNNNNNNNN 1980
QY 2313 GATGAGATCAACAACCTGCGGTGAGTTGACAGTGTGTGAAATGGGCACTTCAACAT 2372
DB 1981 NNN 2040
QY 2373 TGACACCCAGCTCAGAAATAAGTTCACTCTCAAGTGTGAGCCATGCTGAGAGCTG 2432
DB 2041 NNN 2100
QY 2433 CGGCTGTGCTCAAGGCTGACCCAGACTTCCGATGTGCTGTGCGCAGGCGCCAGGCGCA 2492
DB 2101 NNN 2160
QY 2493 GTGCACTTGGCGCAGACTGCGCTGCCAGAGAGCAAGTGCCTGAGAGTGTGTGTGTC 2552
DB 2161 NNN 2220
QY 2553 CAAAAGCAGTGCACAAAACCCCGCATCACAGAGATAATCCGATGACAGGCCCGCGGA 2612
DB 2221 NNN 2280
QY 2613 AGGGGAGCACAAGGTCACTATCCGAGGGAGAACTGGGCGCTGGAATTCGCGACATCCG 2672
DB 2281 NNN 2340
QY 2673 CTCCTCAGTCAAGTGTCTGTGCGGTGAGTGAAGCCCTTAGTGAATGATCCCTGC 2732
DB 2341 NNN 2400
QY 2733 AGAAGCAGTGTGTGAGATGGGGGAGGCCAAGCCC--AGCGAGCATGCAAGCTTCTGT 2789
DB 2401 NNN 2460
QY 2790 GGAAGTCTGCGTGTGTGTGTGCGCTGAATTTATGGCCCGGTCTTCAAGCTTAATTA 2849
DB 2461 NNN 2520
QY 2850 CTTATGACACTGACTCTCTCAGATCTGAAGCCAGCGGGGCGCCATGTCCGAGGGAGC 2909
DB 2521 NNN 2580
QY 2910 CCAAGTGAATCAACAGCACAACCTGAATGCGGAGAGCAAGTGTGTGTGTGTG 2969
DB 2581 CTGATTTGGCAATGAGGAGCACTGAAGCAGGAGATGATGTGCTGTGTGTGCGG 2640
QY 2970 AAAGAGCCCTGTCTCTTCCACAGCGCATCTCACTCATGTGTGCAACCAATC 3029
DB 2641 TGGCGGCGCTGTCTCTTCTCTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2700
QY 3030 CTCAGATGAGGTGCTAGAGATGAAGGTGTGCTGAGTGAAGGCGCAAGATCCACCA 3089
DB 2701 NNN 2760
QY 3090 GGAAGTGTCTTTGATGTGAAGACCCACATGTGTGAGTGAAGCAGAAATGAG 3149
DB 2761 NNN 2820
QY 3150 CATGTCAAGTGAACAACACCATGCGGTATGGGGGACCACTGACCTCATACAGAA 3209
DB 2821 NNN 2880
QY 3210 CCCCCAGATCCGTGCAAGCATGAGGAGAGAGACATCAATATCTGTGAGTTCTGAA 3269
DB 2881 NNN 2940

QY 3270 CGTACTGAGATGACCTGTGACAGCGCCGCGCTGTGAGTCTTGACCA--CAAGTC 3326
DB 2941 NNN 3000
QY 3327 AGACCTGACCGAGAGGCGCGAGAGTGTGCTTATCTTGAGACAACCTCAAGTCTGTCT 3386
DB 3001 AGAGTGGGAGAGCGCGAGTGAAGTGTGCTGTGATGACAAATGTGCGCTCTGTCT 3060
QY 3387 CATCTGAACAAGCAACCTTCACTCTATATCCCAACCGGTGTGAGGCTTGTGTC 3446
DB 3061 TGTGCTCACTCACT 3120
QY 3447 CTCAGAAATCTGTGAGTCAAGCTGACCGCCATCATCTTAAGGCGAAACCTGAT 3506
DB 3121 CACTGGCTGTGAGGTGAAGCCAGCTCCCACTCATCTCAAGGGCGGAACTCTT 3180
QY 3507 CCGGCTGTGCTGGGGGCAACGTGAAGCTGAATCACTGTGTGTGTTGGGAGAGCC 3566
DB 3181 GCCACTGTCACT--GGCACTCCGACTCACTCACTCACTCACTCACTCACTCACTCACT 3237
QY 3567 GTGCACTGTGACCGTGTGAGATGTCAGCTGCTGTGAGAGTCCCAACCTCATGCGAG 3626
DB 3238 CTGTACTCTCACTGTGTGAGAGAGCAATGTGTGTGAGAGGCGCCAACTCATGCGGNN 3297
QY 3627 GCAAAAGTGAATGGCGGTGTGCTGTGAGTATCTCCCGGGAGTGTGTATGTC 3686
DB 3298 NCACAAAGTCAAGGTGCGGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 3357
QY 3687 CCGGAGAGCGCGCTGAGCTGCGCGGCACTGTGAGATGCACTGTGCTGCGCTCTCT 3746
DB 3358 CTCGAGAGCTGTGTGAGAGT 3417
QY 3747 CATCATTTTCACTGTGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3806
DB 3418 GGT 3477
QY 3807 CACGCTGAAGGGGTGTGAGATGAGATGAGACAACCTGAGTCCGCTGTGCGCTGAGATG 3866
DB 3478 CACACTCAAGGCGGTGTGAGATGAGATGAGACAACCTGAGTCCGCTGTGCGCTGAGATG 3537
QY 3867 CAAGAGCTTGTGCGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGT 3926
DB 3538 CAANNNNCTTTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGT 3597
QY 3927 AGCGGAGTTCGCTTCTGTGAGTATGAATTAACAATGAGGAGTGTGAGAGTGTGAGAGT 3986
DB 3598 TGGCGGAGTTCGCTTCTGTGAGTATGAATTAACAATGAGGAGTGTGAGAGTGTGAGAGT 3657
QY 3987 TGAAGACCACTGTCTCCGAGACTTGAAGTCCCGGCTTACCGGAGAGCGTGTGGA 4046
DB 3658 CGAGGACCACTGTGTCTCAAGAGATGAG-----NNNNNNNNNNNNNNNNNN 3705
QY 4047 GAAAGCTGAAGTCTTGTGCGGAGCTATCAACAAGTGTGTGTGTGTGTGTGTGTGT 4106
DB 3706 NNN 3765
QY 4107 CCGCAGCTTGAGTCCAGCGTAGTCTTCATATGCGGACCGTGTGCAAGTGTGCTCACT 4166
DB 3766 NNN 3825
QY 4167 CATCATGACCGTGTGAGAGCAAGCTGAGTACCACTGATGTGTGAAGAGAGTGTCT 4226
DB 3826 NNN 3885
QY 4227 GGGCGACCTCATTTGACAGAACTTGAGAGCAAGAACACCTAAGTGTGTGTGAGAG 4286
DB 3886 NNN 3945
QY 4287 GACTGAGTGTGCTGAGAGATGCTGAATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4346
DB 3946 GACTGAGTGTGCTGAGAGATGCTGAATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4005
QY 4347 CCTCAAGAGTGTGCTGAGAGCGCCCTGTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGT 4406

Db	4006	CTTCAAGAGATGCGCTGGGAGCGCGTTCCTACGTCTGATCTGGCCCATCAAGACAGAT	4065
Qy	4407	GGAGAAAGGCCCCATTGACGCGCATACGGCGAGGCCCCGTACTCTTGAGCGAGCA	4466
Db	4066	GGAAAGGCCCCCATTTGACGCGCATACGGGGAAGCAGCTACCTCTGATGAGACAA	4125
Qy	4467	GCTGATCCGACAGAGATTGACTCAAAACCCTGGTCTGAGCGTGTCAAGCCAGACAA	4526
Db	4126	GCTCATCCGACAGAGATTGACTCAAAACNNNNAACCTGAACGTGTGAACCTGTGAA	4185
Qy	4527	TGCCAAGACCCCGAGTCCCATGTAAGATCTCTCACTGTGACCACTCATCTAGTCCA	4586
Db	4186	TGAGATATCATCTGGGCGCGGTGAAGGGGCTGAGCTGTGACAGGTGTACCAAGGCAA	4245
Qy	4587	GGAAGAATTGGATGTCATCTTCAAGATATGTCCTTGCTCCACGGGCCAAAGCTGC	4646
Db	4246	GGAGAAGTGTGGAACGTGCTTCAAGAGCGTGCCTCANNNNNNNNNNNNNNNNNNNN	4305
Qy	4647	AGATATGATCTGAGTGGCGACAGGAAGTGGGGCAAGATGATCTTGACGATGAAGA	4706
Db	4306	NN	4365
Qy	4707	CATCACCAACAAGATTGAATGATTTGGAACGACTGAACAACATGGCCCACTACAGGT	4766
Db	4366	NN	4425
Qy	4767	GCCAGATGTTCCGTGTGGGATAGTGTCCAAAGAGTGAACAGCTATAACGAGTGAA	4826
Db	4426	GACAGACGGGTCTCGGTGGCACTGTGTCCAAACAGACATCCGCTCAATATCTTCAA	4485
Qy	4827	CAACTCCACCGTCTCCAGAGCCTCAGCAAGTAAATATGAACAATGATCCGCTTCAACGGG	4886
Db	4486	CTCTTCCACCTTCAACCAAGTCCCTC - - AGCAGATACAGAGCATGTCTGCCACGGCAG	4542
Qy	4887	CAGCCCCAAGCCTCCGCTCAGCGACACCTATGATCACTCTGAACCTGAGAGTGAAGT	4946
Db	4543	NNNNNNCCGACAGCTGCGCTCGCGCACGCTATGATCAGGCCGACCTGGAGAGCGGCAC	4602
Qy	4947	CAAGATGTGGCACTTATGTGAAGAACCAAGACACGGAACAAGAAAGAGGGGGACCGGGG	5006
Db	4603	CAAGTGTGGCACTGTGTGAAGAACCAAGACACCTGGAACAAGGTGAGGTGACCGCG	4662
Qy	5007	GAGCAAGATGGTGTGGAATCTACCTGACCCGACTCTGGCCAATTAAAGGCACACTGCA	5066
Db	4663	CAGCAAGATGGTCTGGAGATCTACTGACACGGCTTACGGCACCAAGNNNNNNNNNN	4722
Qy	5067	GAAATTGTGATGACCTCTTGAGACCATCTTGACACGGCACACGGTGGCTCTGCGCT	5126
Db	4722	NN	4782
Qy	5127	GCCCCGGCATCAAGTACATGTTTGACTTCTGTGATGAGACAGCGTGAATAACATGGCAT	5186
Db	4783	GCGGTGGNATCANNNNCACTGTTGACTTCTGTGATGNNNNNNNNNNNNNNNNNNNN	4842
Qy	5187	TGATGACCCGCACTGTCGCGCATACTCGGAAGAGCAATTGCTGCCCCGAGGTTTGGCT	5246
Db	4843	NN	4902
Qy	5247	CAACATGATCAAGAACCCGCAAGTTGTGTTGATCATCAATGAAGACAGATCAAGACGC	5306
Db	4903	GAACTGATCAAGAACCCCAAGTTGTGTTGATCATCAAGAACAGATCAAGACGC	4962
Qy	5307	CTGCTCTCTGTGTGTGGCTCAAGCTTCAATGACATCTTGCTCCAAGTCAAGACCGGCT	5366
Db	4963	CTGCTGTGTGTGTGGCCCAAGCTTCAATGNNNNNNNNNNNNNNNNNNNNNNNNNN	5022
Qy	5367	GGGCAAGACATCGCCTCCAAAGAGTGTGTATGCAAGACATCCCAAGTACAAAGAA	5426
Db	5023	NN	5082
Qy	5427	TTGGGTGGAGAGGTTATCTCAGACTAGGAAGATGCCAGGCACTCAGGACCAAGACAT	5486

[illegible]

Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koya, S., Kurinara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.

Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sec.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers
1. 3355
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:C330023B09"
/db_xref="WGI:2416574"
/db_xref="taxon:10090"
/clone="C330023B09"
/cell_type="ES cells"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
1. 2210
/note="unlabeled protein product; PLEKXIN 3 homolog [Mus musculus] (SPTK1P70208, evidence: PASTV, 99.7%id, 39.2%length, match=2205) putative"

FEATURES
source
Location/Qualifiers
1. 3355
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:C330023B09"
/db_xref="WGI:2416574"
/db_xref="taxon:10090"
/clone="C330023B09"
/cell_type="ES cells"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
1. 2210
/note="unlabeled protein product; PLEKXIN 3 homolog [Mus musculus] (SPTK1P70208, evidence: PASTV, 99.7%id, 39.2%length, match=2205) putative"

CDS
/protein_id="BAC33681.1"
/db_xref="GI:26340036"
/translation="DYKPGSHVYLKKNLIPAAAGSRINVTYVILGGPCALTVSDTO
LIDDSBQGRQPMVNLVSGLEFWLGLHTADRALTPAMGLAAGGLILATTV
LVAYKRTQADRLKRLQOMDNLSRFAEKEKFAELQDINLNTMGVOJPE
LDRTYAVRVLPGIEAHVPLKEIDLPVNEKALRLFGQLHSRAFLFLFHTLEAS
SFGMRDGTVAALVVALGRLDVALGTLQIADLIEKLSKHPKILIRTSIVA
EKLTNMFPLHLKPLKEKAGEPLFLYCAIKOOMEKGLDITGARYLSBEDKIR
QCIDYKTLTLHCVCPESESAOYPAVYLNDSITOKXKQIDITGARYLSBEDKIR
MDLBMROGRARILIDQEDITTKIECDKRVNLSIAVYQDLSIALVAPKOYAYMA
NSLFTSRISRYESLLNPASSPDLKSRAPMLTPDQAGAKIMHYRNPDHDEGD
RGSQWSEIYLRLATKGLKFDVDEFTVSTHRSALPLAKINFDLDEAD
QKQISPDVHTWKSNCPLRFWVNIKNQPFVDIHKNSITDCLSVAAQTFMDSCS
TSEHRLGKSDSPSKLYAKDIYNTKSWERYDIARMSISPDQDAVLVQSRILHA
NDPVLVSLSELVYVTKYRGELITSLDRDASCRKHLKQKLEQIITLVSSSS"

polya_signal
3332..3337
/note="putative"
3355
polya_site
/note="putative"

ORIGIN
Query Match 21.5%; Score 1221.2; DB 11; Length 3355;
Best Local Similarity 72.8%; Pred. No. 1,4-246;
Matches 1622; Conservative 0; Mismatches 588; Indels 18; Gaps 3;
QY 3458 TGAAGCTCAAGCTGGCAAGCCCATCACTAAAGGCAAGAACTTATCCCGCTGTG 3517
Db 1 TAGATGCAACCTGGTTCACATGTTGATGAAGGCAAGAACTTATCCCGCTGTG 60
QY 3518 CTGGGGGCAAGCTGAAGCTGCTGCTGCTGGGGAAGAACCCGTGACCGTGA 3577

Db 61 CTGGCAAGCTCCCGC---CTCACTACCAAGTGTGGATTGGAGACAGCACTGTGCACTAA 117
QY 3578 CCGGTGAGATGTCAGACTCTCTGCGAGTCCCAACTCATCGCAGGCAAAAGTGA 3637
Db 118 CTGCTCAGATCACTCAACTTCTGTGTATCCCAAGCAGACAGCGCGCAAGCTGTTA 177
QY 3638 TGGCCGTGCTGGTGGATGAGTACTCCCCGGGATGTGTATTCGCCCCGGAAGCC 3697
Db 178 TGGGTGCTGGTGGTGGCTGGAGTTCGTGTGGGCAAGCTGACATCACTGCTGATCGGG 237
QY 3698 CGCTCAGCCCGCCCGCATGTCAGATGAGATGAGTGGCTGGCGCTCTCATATTTCA 3757
Db 238 CACTGACCTTACCAAGCTATGTGTGGGCTTACAGACAGGCGCGCTTGTCTGTGCGCA 297
QY 3758 TCGTGGCCGTGCTCATTTGCTTATTAACGAAATCCGCGAAAGTACCTGACCTGAAGC 3817
Db 298 TCACTGTTGTGCTGTGCTTACCAAGGCTAAGCTCAGATGAGAACCGCACCTTAAGC 357
QY 3818 GCGTCAGATGAGATGAGAACCTGAGATCCGCTGTGGCCCTGAGATGAGAAAGCCT 3877
Db 358 GCGTTCAGCTACCAATGAGACACTGAGATCTGTGTGGCTTGGATGAGTGAAGAGCCT 417
QY 3878 TTGCGAGCTGCAAGCCGACATCATGAGTGAACCACTGATGAGACCGGAGTTC 3937
Db 418 TTGCTGAGCTGCAAGCTGATATGATGATGATGATGATGATGATGATGATGATGATG 477
QY 3938 CATTCTGAGCTATGAACTTACCACTGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 3997
Db 478 CATTCTGAGCTATGAACTTACCACTGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 537
QY 3998 CTGCTCTCCGGGACCTTGAAGTCCCGGAGTACCGGCGAGAGCGTGTGAGAAAGCCTGA 4057
Db 538 CAGTGTCTGAGAGTCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 585
QY 4058 AGCTCTTGGCCAGCTCATCAACAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4117
Db 586 GTCTATTGAGCAAGTCTGCTGCAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 645
QY 4118 AGTCCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4177
Db 646 AGGCTCAGAGTGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 705
QY 4178 TGTGCAAGAGCAAGCTGAGTACCCACTGATGTGTGTAAGACAGCTGTGGCGCACTCA 4237
Db 706 CCTGCAAGAGCGGCTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 765
QY 4238 TTGCAAGAACTGAGAGGCAAGAACCACTGAGCTGCTGCTGAGAGGAGTACAGTACAG 4297
Db 766 TAGAGAAAACCTTGAAGAGCAAGAACCACTGAGCTGCTGCTGAGAGGAGTACAGTACAG 825
QY 4298 TGGCTGAGAGATGCTGCAATTTGTTACTTCTCTCTCTCAAGTTCTCAAGAGT 4357
Db 826 TGGCTGAGAGATGCTGCAATTTGTTACTTCTCTCTCTCAAGTTCTCTCAAGAGT 885
QY 4358 GTGCTGGGAGACCT 4417
Db 886 GGGCGGGAGCACT 945
QY 4418 CCATGACGCAATCAAGGCGGAGGCGGCTACTCTCTTGAAGGAGACAAAGCTCATTCGCGC 4477
Db 946 CTAATGAGCAATTAAGGCGGAGGCGGCTACTCTCTTGAAGGAGACAAAGCTCATTCGCGC 1005
QY 4478 AGCAATGATCACTAAACCTCTGCTGAGCTGTGTAGCGCGCAAGATGCAAGC 4537
Db 1006 AGCAATGATCACTAAACCTCTGAGCTGTGTAGCGCGCAAGATGCAAGC 1065
QY 4538 CCGAGTCCAGTAAAGATCTCACTGATGACACATCACTCAAGTCAAGAGAAAGATTC 4597
Db 1066 CCGAGTCCAGTAAAGATCTCACTGATGATGATGATGATGATGATGATGATGATGATGATG 1125
QY 4598 TGAATGCACTTCAAGATGTGCT 4657

Db 1166 TGGATCTGTGTACAAAGGATATTCTCAGCGCCCAAAAGCTGAGACATGGATT 1185
 Qy 4658 TGGATGGGCAAGAAAGTGGGCAAGATGATCTTGCAGGATGAAACATCAACCA 4717
 Db 1186 TGGATGGGCAAGGCGCGCATGCGCCGATCATCTCCAGATAGAGACATCAACAA 1245
 Qy 4718 AGATTGAGATGATGAAAGCAGCTGAAACACCTGGCCCACTACCAAGGTGCCAATGGTT 4777
 Db 1246 AGATTGAGTGAAGTGAAGAGAGGTCAACTATTGGCCCACTACCAAGGTGCTATGGTT 1305
 Qy 4778 CCGTGGTGGCTATTAGTGTCCAGAGAGTGAAGCTTATAGCAAGTGAACAACTCAACG 4837
 Db 1306 CTTTATGATGACCTGTGTGCGCAAAAGAGTGTCTGCTATTAATGAGCCAACTGTTACCT 1365
 Qy 4838 TCTCCAGAGCTGAGCAAGTAAATATGAAAAATATATCCGGTACACGGGCGAGCCGACCA 4897
 Db 1366 TCAACCGG---TCACTTATGCTGCTACGAGAGCTTCTCCGTGCTCCAGCGCCGAGTA 1422
 Qy 4898 GCGTCCGCTCAGGAGACCTATGATCACTGACCTGAGAGTGAAGTCAAGATGTGC 4957
 Db 1423 GCGTCCGCTCCGAGACCTATGCTACGCTGACCGCCGACAGAGGAGGTGCCAAGCTGTGC 1482
 Qy 4958 ACCTAGTGAAGAACCGAGACCGAGACCGAGAGAGGAGGAGCCGAGGAGCAAGATG 5017
 Db 1483 ACCTGGTGAAGAACCGAGACCGAGACCGAGAGAGGAGAGCCGCGAGCAAGATG 1542
 Qy 5018 TGTCTGAATCTAACCTGACCGGACTCTGCGCCCACTAAGGGGCACTGAGAGAGTTGTG 5077
 Db 1543 TGTCAAAATATATCTCAAGAGGCTGCTGCGCCCAAGGGGCACTGAGAGAGTTGTG 1602
 Qy 5078 ATGACCTCTTTGAGACCATCTTCAAGACGCAACCGGTGCTGTGCTGCTGCTGCGCA 5137
 Db 1603 ATGACTGTTTGAATCTGTGTGATGACAGCCGAGCGGAGCTCAAGCTTACCTTGGGCA 1662
 Qy 5138 TCAAGTACATGTTGATCTTCTGAGAGAGAGGCTGATTAAGTGGCATTCATGACCGG 5197
 Db 1663 TCAAGTACATGTTGATCTTCTGAGAGAGAGGCTGATTAAGTGGCATTCATGACCGG 1722
 Qy 5198 ACCTCGGCATACCTGGAAGAGCAATGCTGCGCCCTGAGGTTTGGGTCAACATGATCA 5257
 Db 1723 ATGTGGTCACTGCTGGAAGAGCACTGCTTACCTGCTTCTGCGGTGATGATGATCA 1782
 Qy 5258 AGAACCGGCACTTGTGTGATGATCAATCAAGAGCAATGATCAAGAGCGCTGCTG 5317
 Db 1783 AGAATCGCAATTTGTGTGATGATCAATCAAGAGCAATGATCAAGAGCGCTGCTG 1842
 Qy 5318 TGTGGGCTCAGACCTTCATGATCTTGTGCTCCAGCTCAGAGCAACCGGCTGGGCAAGACT 5377
 Db 1843 TGTGGGCTCAGACCTTCATGATCTTGTGCTCCAGCTCAGAGCAACCGGCTGGGCAAGACT 1902
 Qy 5378 GCGGCTTCAACAGCTGCTGATGATCAAGAGCAATCCCGAGCTTCAAGAAATTGGGTGAGA 5437
 Db 1903 CACCTTCAACAGCTGCTGATGATCAAGAGCAATCCCGAGCTTCAAGAAATTGGGTGAGA 1962
 Qy 5438 GGTATTCTCAGACATGAGAGAGATGAGAGCACTCAGAGCAACGAGCATGAACGATACC 5497
 Db 1963 GGTATTCTCAGACATGAGAGAGATGAGAGCACTCAGAGCAACGAGCATGAACGATACC 2022
 Qy 5498 TGGCTGAGAGTCCCGAGTCAATGATGATGATCAACAGATGATGATGATGATGATGATG 5557
 Db 2023 TGGCTGAGAGTCCCGAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 2082
 Qy 5558 TCTTCTCTATGAGAGAGATGAGAGAGATGAGAGAGATGAGAGAGATGAGAGAGATGAGAG 5617
 Db 2083 TCTTCTCTATGAGAGAGATGAGAGAGATGAGAGAGATGAGAGAGATGAGAGAGATGAGAG 2142
 Qy 5618 AGTGTGGAAGAGAGAACTGGCTTCAACAACTGAACAACTGAACAACTGAACAACTGAAC 5677
 Db 2143 CTGTGTGGAAGAGAGAACTGGCTTCAACAACTGAACAACTGAACAACTGAACAACTGAAC 2202
 Qy 5678 ACAGCTGA 5685
 Db 2203 GCAGCTGA 2210

RESULT 9
 CD653539
 LOCUS
 DEFINITION
 AGENCOURT 14542448 NIA Human H1 Embryonic Stem Cell cDNA Library
 (Long) Homo sapiens cDNA clone IMAGE:30420107 5', mRNA sequence.
 CD653539
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 838)
 NIH-MGC http://mgs.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: dsgerha@mail.nih.gov
 Tissue Procurement: Irene Gintis and Mahendra Rao, NIA
 cDNA Library Preparation: Yulan Piao and Minoru Ko
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINTL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC c
 can be found through the I.M.A.G.E. Consortium/LINTL at:
 http://image.llnl.gov
 Plate: NDMA95 row: m column: 12
 High quality sequence stop: 650.
 Location/Qualifiers
 1..838
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30420107"
 /issue_type="Embryonic Stem cells"
 /cell_line="WA01"
 /lab_host="DH10B (TI phase-resistant)"
 /clone_idb="NIA Human H1 Embryonic Stem Cell cDNA Library (Long)"
 /note="Vector: PCMV-Sport6; Site 1: NotI; Site 2: SalI;
 This is a long-transcript enriched cDNA library (Genome
 Res. 11: 1553-1558 (2001) [PMID: 11544199]) from WA01/
 cell line. Undifferentiated human ES cell line WA01/H1
 was obtained from WiCell Research Institute, Inc.,
 Madison, WI, cultured according to their instructions, on
 MEF feeders. They formed round colonies with defined edges
 and were positive for alkaline phosphatase, SSEA-4, OCT3,
 OCT4, REX1, UTR, TERT, SOX2, CX43 and CX45. They are
 negative for GAP2, GAP4, PDX1, NCM, MSX1, FLT3, SSEA-1,
 TUBB3, NES, GFAP, and EOMES. When colonies (18-10 days
 after plating), the ES cells from 4 x 6cm dishes were
 treated with 1 mg/ml collagenase, type IV
 (Invitrogen/GIBCO) for 5-10 min and gently scraped off
 with 5 ml pipette. RNA was purified with TRIzol Reagent
 from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558
 (2001) [PMID:11544199] Double-stranded cDNAs were
 synthesized with an oligo(dT) primer [Invitrogen:
 5'-pGATTAAGTCTAGATCCGAGCGGCGCCCTTTTCTTTT-3'] from
 3.4g of total RNA, treated with 14 DNA polymerase, and
 purified by ethanol-precipitation. The cDNAs were ligated
 to lone-linker lp-Sal4, purified by phenol/chloroform
 extraction, and separated from free linkers by
 centrifugation on a 100 column. Then, the cDNAs were amplified by
 long-range high fidelity PCR using Ex Tag polymerase
 (Takara) with a primer Sal4-S for 25 cycles. The products
 were purified by phenol/chloroform extraction and
 centrifugation on a 100 column. The cDNAs were digested with SalI
 and NotI enzymes and cloned into SalI/NotI site of
 PCMV-Sport6 plasmid vector. The average insert size is

QY	4244	CAGAAGCTCGAGACAGAAACACACCCTAACCTCTCTCTCAGAGAGACTGAGTCAGTGGC	4301
Db	491	TAAAGAACTCGAGAAACAAAGAACACCCCAAGCTCTACTCCGAGAGACAGAGTCTGTGGC	550
QY	4302	TGAGAAAGATGCTGACCAATTGGTTTACTTTCCTCTCTTACAGTTCTCTCAAGAGATGTGC	4361
Db	551	TGAAAAAGATGCTGACCAATTGGTTTGCCTTCTCTCTCTCAAGTTCTTAAAGAGTGTGC	610
QY	4362	TGGAGAGCCCTCTTCTCCCTGTTCTGTGCATCAAGCAGACAGATGAGAGAGGCCCAT	4421
Db	611	AGGGAGGCCACTCTTCTATGCTATCTGTGCGCATTAAGCAGACAGATGAGAGAGGCCCAT	670
QY	4422	TGAGCCCATACGGGCGAGGCCCGCTACTCTTGTAGCGAGAGACAGCTCATTCGCGAGA	4481
Db	671	TGATGCACTACGGGCGAGGCCCGCTACTCTCTGTAGGAGAGACAGCTCATTCGCGAGA	730
QY	4482	GATGACTCAAAAACCTGCTGCTGTAGCTGTGAGCGCCAGACATGTCCACAGGCCCGCA	4541
Db	731	GATCGAGTCAAAAGACCTGTATCTTCTTAATCTGTGATCACTTGCAGACAGAGAACATGTCCAG	790
QY	4542	GGTCCAGTAAAGATCTCTCAACTGTGACACCATCACTCAAGTCAAGAGAGAAATTTGGA	4601
Db	791	GATCCCAAGTAAAGTGTAACTGTGACACCATCCCCCAAGTCAAGAGAGAAATTTTGA	850
QY	4602	TGCGACTTTCAGAAATGT-GCCTTGCTCCGACCGGGCCAAAGCTGACAGATATGAAATCTGG	4661
Db	851	TGCGGGGATTAAGAAATGTGGCCCTTATTTCCACGCGCGAGGGGCAGTGACATGGACTTGG	910
QY	4661	AGTGGCGACAAAGG-AAAGTGGGCAAGATGATCTTGCAGATGAAGAACATCACCA	4716
Db	911	GATGGCGGCAAGGCCCGGATGTGCGGGCTTGAGATGAGACATCCCCCCCCA	970
QY	4719	GATTGAGA 4726	
Db	971	GATTGAGA 978	

RESULT 11	AL529477	1201 bp	mRNA	linear	EST 23-MAY-2000
LOCUS	AL529477				
DEFINITION	AL529477 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens				
	CNA clone CS0DD006y12 5-PRIME, mRNA sequence.				

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eumetazoa, Metazoa; Chordata; Crustacea; Vertebrata; Butelosostrum;
Mammalia; Eutheria; Primates; Carnivora; Homnidae; Homo.
1. 1 bases 1 to 1201
L. M. B., Gruber, C., Jessup, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 13, 2001 this sequence version replaced gi:12792970.

FEATURES
source

```
1. 1201
   /organism="Homo sapiens"
   /mol_type="rRNA"
   /db_xref="taxon:9606"
   /clone="CSDD006YE12"
   tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"
```

```

/clone lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-1000(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

```

Query Match	11.4%;	Score 650.6;	DB 9;	Length 1201;
Best Local Similarity	82.5%;	Pred. No. 3e-126;		
Matches 780;	Conservative 1;	Mismatches 160;	Indels 4;	Gaps 3

QY	377	TTGGCTATTAAGCGCAAGTCCCGCGAAATGACCTCAGCGTGAACGGCTCAGATGAGA	3833
Db	63	TTGCTCTCAAGCGCAAGTCTCGAAGAAATGACTCTCTCAACGGCTCGAATGAGA	1222
QY	3833	TGGACAACCTGGAGTCCCGTGTGTCCTCGAGTGCAGAAAGACTTTGCCAGCTGCAGA	3888
Db	123	TGGACATCTGGAGTCCCGTGTGTCCTCGAGTGCAGAAAGCTTTTGTGAGTCTCAGA	1822
QY	3893	CGGACATTCATAGCTGACCGAGTGCCTGGATGGAGCCGGGATTCGCTCTCGACATTA	3953
Db	183	CGGATATTCATAGATTGACAGTGTGACCTTGGACCGCTCAGGAATCCCTTACTGGAATTC	2424
QY	3953	GAACTTACACCATGCGGGTGTCTGTTCCAGGAATTGAAGCAACCTGTCTCCGGAGCC	4013
Db	243	GTACTAGCTATGGAGTCCGTTCCGGGGATGAGAGCAACCCGTCTGTGGGAGGC	3023
QY	4013	TTGAGTCCCGGGCTTACCGGCGAGAGCGTGTGGAGAAAGGCTGAAGCTTTGGCCGAC	4073
Db	303	TGGAGGTAAAGAGAAACGGGCGAGAGCAGTGGAGAAAGCCCTGAAGTCTTTGCCAGC	3623
QY	4073	TCATCAACAACAAGGTTCTCTGTGCTTCATCCGACCGCTTGAGTCCAGCGTAGCT	4133
Db	363	TCATCAACAACAAGGTTCTCTGTGACCTTCATCCGACCCCTGAGAGTGCAGCGATT	4222
QY	4133	TCTCCATGCGGACCGTGGCAACGTGGCTCATCTATATGACCGTGTGTGAGAGCAAGC	4193
Db	423	TCTCCATGCGGACCGGCGCAACGTGGTCTGTGCTCATATGACCGGCTTGAAGGAC-SCC	4813
QY	4193	TGGAGTACGCACTGATGTGTGGAAGCAGCTGTGGCGCATTCATGACAAAGAACCTGG	4253
Db	482	TGGAATATGTCACTGATGTGTCTCAAGCAGCTCTCTGACTCATATGATTAAGAACTGG	5423
QY	4253	AGAGCAAGAACCAACCTTAAGTGTCTTGAAGAGACTGAGTCAGTGTGAGAAATGC	5423
Db	542	AGAACCAAGAACCAACCCAAAGCTGTACTTCGGAGAGACAGAGTCTGTGTGAAAAAGATGC	6013
QY	4313	TGACCAATTGATTATCTTCCCTCTCTCAAGTCTCTAAGAGATGAGAGATGAGAGAGCCCC	4313
Db	602	TGACCAATTGATTGTGCTCTCTCTCTCTGACAAATTTCTTAAGAGATGCGCAGGAGGCCAC	6623
QY	4373	TCTTCTCCCTGTTCTGTGTGCATCAAGCAGACAGATGAGAGGGCCCCATTGACGCATCA	4473
Db	662	TCTTCAATGCTATCTGTGCATCAAGCAGACAGATGAGAGGGCCCCATTGATGCATCA	7213
QY	4433	CGGCGCAGAGCCCGCTACTCTCTTGACGAGAGCAAGCTCATCCGCGAGAGATTGACATCA	4453
Db	722	CGGCGCAGAG-CGCGTACTCTCTTGACGAGAGCAAGCTCATCCGCGAGAGATTGAGATCA	7823
QY	4493	AAACCTGTGCTGAGCTGTCAAGCCGACGAACAATGCCAGACGCCCGAGGTCTCCAGTTA	4553
Db	781	AGACCTGATCTCGAATGTGCTCAACCTGACAAAGAGAACGTCCAGAGATCCAGATGA	8443
QY	4553	AGATCTCAACTGTGACACCATCATCTCAAGTTCMAAGAGAAATTTCTGATGCCATCTCA	4613
Db	841	AGATGTTAAACTGTGACACCATCAACAGGTCAAGAGAAAGTTTGTATGCGGTATTA	9013
QY	4613	AGATATGTGCTGTGCTCCACCGGCGCAAGCTGACAGATATGATCTGGAGTGTGCGACAAG	4673
Db	901	AGATATGTGCTATTTCTCCAGCGCGCGAGGCGCAGTGAATGGAATTGGAGTGTGCGACAAG	9613
QY	4673	GAAATGGGAGAGATGATCTTGCAGAGATGAAGATCAACCA 4117	

Db 961 CCGATCGCCGGTCTGCT--GCATGAGACATCACCA 1003

RESULT 12
CA316867 745 bp mRNA linear EST 09-JUL-2003
LOCUS UI-M-FW0-cbk-j-22-0-UI-r1 NIH_BMAP_FW0 Mus musculus cDNA clone
CA316867 IMAGE:6810839 5', mRNA sequence.
DEFINITION CA316867.1 GI:24534991
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
MUS musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 745)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pyx-5.
FEATURES
source
1. 745
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6810839"
/tissue_type="whole brain"
/dev_stage="embryo 13.5, 14.5, 16.5, 17.5dpc"
/lab_pos="DHI08 (T1 phage resistant)"
/clone_1lb="NIH_BMAP_FW0"
/note="Organ: Brain; Vector: pyx-Asc, Site_1: Ecor I;
Site_2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with Ecor I adaptor, digested with NotI and then cloned
directionally into pyx-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGCGAGACAG. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemlin Chitt, Ph.D.,
program coordinator."

ORIGIN
Query Match 11.2%; Score 638.4; DB 14; Length 745;
Best Local Similarity 91.0%; Pred. No. 8.7e-124;
Matches 678; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

Db 1 CAAGGAGCTTTGCGAGCTGCAGCAATATTCATGCTTACCAAGTCACTGATGG 60

QY 3867 CAAGGAGCTTTGCGAGCTGCAGCAATATTCATGCTTACCAAGTCACTGATGG 3926

Db 3927 AGCGGAGATTCGCTTCCTGACTATAGAACTTACACCATGCGGGTCTTTCCAGGAT 3986

Db 61 AGCTGGATTCCTTTCTGCACTACAGAACTATACCAAGCGGGTCTTTCCAGGAT 120

QY 3987 TGAAGACCACTGTCTCTCCGGACCTTGAGTCTCCGGCTACCGGACGAGCGTGTGA 4046

Db 121 TGAAGACCACTGTCTCTCCGGACCTTGAGTCTCCGGCTACCGGACGAGCGTGTGA 180

QY 4047 GAAAGCCTGAAAGCTCTTGCCCGAGCTCATCAACAAGAGTGTCTGCTGCTTCAT 4106

Db 181 GAAAGCCTGAAAGCTCTTGCCCGAGCTCATCAACAAGAGTGTCTGCTGCTTCAT 240

QY 4107 CCGCAGCTTGAGTCTCCAGAGTGTCTTCCATGCGCCGACCGTGGACAGTGGCCCT 4166

Db 241 CCGCAGCTTGAGTCTCCAGAGTGTCTTCCATGCGCCGACCGTGGACAGTGGCCCT 300

QY 4167 CATCATGACCTGTCTGCTGAGCAAGCAAGTGTGACCAATGTGCTGAGAGCTGTCT 4226

Db 301 CATCATGACCTGTGCTGAGCAAGTGTGACCAATGTGCTGAGAGCTGTCTGAG 360

QY 4227 GCGCAGCTTGAGTGTGAGCAAGTGTGAGCAAGTGTGAGCAAGTGTGAGCAAGT 4286

Db 361 GCGTGAACCTGATGATGAGAACTGTGAGCAAGTGTGAGCAAGTGTGAGCAAGT 420

QY 4287 GACTGATCAGTGTGAGTGTGAGCAAGTGTGAGCAAGTGTGAGCAAGTGTGAGCA 4346

Db 421 GAGGAGTGTGAGTGTGAGCAAGTGTGAGCAAGTGTGAGCAAGTGTGAGCAAGT 480

QY 4347 CTTCAAGAGTGTGAGTGTGAGCAAGTGTGAGCAAGTGTGAGCAAGTGTGAGCA 4406

Db 481 CTTCAAGAGTGTGAGTGTGAGCAAGTGTGAGCAAGTGTGAGCAAGTGTGAGCAAGT 540

QY 4407 GGAAGAGGCGCCGATGAGAGTGTGAGCAAGTGTGAGCAAGTGTGAGCAAGTGTG 4466

Db 541 GGAAGAGGCGCCGATGAGAGTGTGAGCAAGTGTGAGCAAGTGTGAGCAAGTGTG 600

QY 4467 GCTCATCCGCGAGCAAGTGTGAGCAAGTGTGAGCAAGTGTGAGCAAGTGTGAG 4526

Db 601 ACTCATCCGAGCAAGTGTGAGCAAGTGTGAGCAAGTGTGAGCAAGTGTGAGCA 660

QY 4527 TGCCCAAGGCGCCGATGAGAGTGTGAGCAAGTGTGAGCAAGTGTGAGCAAGTGTG 4586

Db 661 TGTTCAGGCGCCGATGAGAGTGTGAGCAAGTGTGAGCAAGTGTGAGCAAGTGTG 720

QY 4587 GGAAGAGTGTGAGTGTGAGCAAGTGTGAGCAAGTGTGAGCAAGTGTGAGCA 4611

Db 721 AGAGAGATCTGATGATGCACTTTC 745

RESULT 13
BX470595 618 bp mRNA linear EST 04-SEP-2003
LOCUS DKEZP686G07110.F1 686 (synonym: hloc3) Homo sapiens cDNA clone
DEFINITION BX470595 mRNA sequence.
ACCESSION BX470595
VERSION BX470595.1 GI:31664922
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 618)
Koehler, K., Beyer, A., Mewes, H.W., Well, B., Amlid, C., Osanger, A.,
Rodo, G., Han, W., and Wiemann, S.
EST (Koehler, K., Beyer, A., Mewes, H.W., Well, B., Amlid, C., et al.)
Unpublished (2003)
CONTACT: MIPS
MIPS
Ingolstaedter Landstr. 1, D-85764 Neuberg, Germany
This is the 5' sequence of the clone insert.
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de;
Sequenced by BMFZ (Biomedical Research Center at the Heinrich-
Heine-University, Dueseldorf/Germany) within the cDNA sequencing
consortium of the German Genome Project. No sl sequence available.
This clone (DKFZP686G07110) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubergweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers

```
1..618
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKZPg686G07110"
/dev_stage="adult"
/lab_host="DH.10B"
/clone_id="686 (synonym: h1cc3)"
/note="Vector: pT7P1Ex2; Site_1: SfiI; Site_2: SfiI;
CDNA-collection"
```

```
Query Match      10.8%; Score 613; DB 13; Length 618;
Best Local Similarity 100.0%; Pred. No. 1.8e-118;
Matches 613; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY	2683	AAGGTCTCGCCGTGGAGTGAAGCCCTTAACTGATGATGTTACATCCCTGCGAACAATC	2742
Db	6	AAGGTCTCGCCGTGGAGTGAAGCCCTTAACTGATGATGTTACATCCCTGCGAACAATC	65
QY	2743	GTCTGTGAGATGGGGGAGGCCAACCACCGACGATGAGAGCTTCGTGGAAATCTGGTG	2802
Db	66	GTGTGTGAGATGGGGGAGGCCAACCACCGACGATGAGAGCTTCGTGGAAATCTGGTG	125
QY	2803	GCTTGTGTGCGCCCTGGAATTCATGAGCCCGGTCTCACAGCTCTATTACTTCATGACATG	2862
Db	126	GCTGTGTGTGCGCCCTGGAATTCATGAGCCCGGTCTCACAGCTCTATTACTTCATGACATG	185
QY	2863	ACTCTTCAGATCTGGAAGCCCAAGCCGGGGGCCAATGTCGGAGGGAGCCCAAGTACATC	2922
Db	186	ACTCTCTCAGATCTGGAAGCCCAAGCCGGGGGCCAATGTCGGAGGGAGCCCAAGTACATC	245
QY	2923	ACAAGGACCAACTGAATGCCGAAGCAACGTGTGTGATGTTTGGAAAGCAAGCCCTGT	2982
Db	246	ACAAGGACCAACTGAATGCCGAAGCAACGTGTGTGATGTTTGGAAAGCAAGCCCTGT	305
QY	2983	CTCTTCCACAGGGCATCTCCATCTCTACATTGTCTGCAACACACATCCTCGATGAGGTG	3042
Db	306	CTCTTCCACAGGGCATCTCTCCATCTCTACATTGTCTGCAACACACATCCTCGATGAGGTG	365
QY	3043	CTAAGATGAAGGTGTGCGTGTGAGGTGGAACAGGGCCAGATTCACACAGGACTGTCTTT	3102
Db	366	CTAAGATGAAGGTGTGCGTGTGAGGTGGAACAGGGCCAGATTCACACAGGACTGTCTTT	425
QY	3103	CAGTATGTGAAGAAGCCCAACCATCTGTGGGATTGAGCCAGAAATGAGCAATGTCACTGGA	3162
Db	426	CAGTATGTGAAGAAGCCCAACCATCTGTGGGATTGAGCCAGAAATGAGCAATGTCACTGGA	485
QY	3163	AAACACCCCATGCGCGTATGGGGGACCACTTGACCTTCATACAGAACCCCAAGTCCGT	3222
Db	486	AAACACCCCATGCGCGTATGGGGGACCACTTGACCTTCATACAGAACCCCAAGTCCGT	545
QY	3223	GCCAAGCATGGAGGAGAGAGACATCAATATCTGTGAGTTCTTGAACGCTACTGAGATG	3282
Db	546	GCCAAGCATGGAGGAGAGAGACATCAATATCTGTGAGTTCTTGAACGCTACTGAGATG	605
QY	3283	ACCTGTCAAGCGCC 3295	
Db	606	ACCTGTCAAGCGCC 618	

RESULT	14				
LOCUS	BM544169				
DEFINITION	BM544169	1027 bp	mRNA	linear	EST 20-FEB-2002
	AGENCOURT_6490555 NIH_MGC_125	Homo sapiens	CDNA	clone IMAGE:5587655	
	5', mRNA sequence.				
ACCESSION	BM544169				
VERSION	BM544169.1	GI:18775199			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Mammalia; Eutheria; Chordata; Vertebrata; Euteleostomi;				
	Eumetazoa; Metazoa; Primates; Carnivora; Homiidae; Homo.				

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
1 (bases 1 to 1027)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC).
Unpublished (1999)
Contact: Robert Straubeberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LM121357 row: 1 column: 02
High quality sequence stop: 676.
Location/Qualifiers
1..1027
/organism="Homo sapiens"

ORIGIN			
Query Match	10.5%	Score 599.2	DB 12; Length 1027;
Best local similarity	78.9%	Pred. No. 2e-115;	
Matches 764; Conservative	0;	Mismatches 139;	Indels 6; Gaps 4;

Db	4280	CCAGAGAGACTGATCAGTGCCTGAGAAAGTGCAGCAATTGGTTACTTCTCTCT	4339
Db	12	TCTCGAGATCTGATCGGTGGCAGAGAAAGTGTACTAACTGGTCACTTCTCTGT	71
Qy	4340	ACAAGTTCTCTAAGAGATGTGCTGGGGAGCCCTCTTCTCCCTGTTTGTGCGCATCAAG	4359
Db	72	ATAAGTCTCTCAAGAGATGGCGCTGGGGAGCCGCTGTTTCATGCTGTACTGTGGCCATCAAG	131
Qy	4400	AGCAGATGGAGAAAGGGCCCATTTAGACGCCATCAACGGCGAGAGCCCGCTACTCTTGGAGC	4459
Db	132	AGCAGATGGAGAAAGGGCCCATTTAGACGCCATCAAGGGTGAAGCAGCGTAACTCCCTGAGTG	191
Qy	4460	AGGACAACTCATCCGCCAGACAGATTGACTTCAAAACCTTGTGTTCTTGAAGCTGTGACGC	4519
Db	192	AGGACAACTCATCCGCCAGACAGATTGACTTCAAAACCTGACCTTAACTGTGTGTAAC	251
Qy	4520	CAGACAAATGCCAACAAGCCCGAGGTCCCAAGTAAAGTCCCACTGTGACACCATCACTC	4579
Db	252	CTGAAATGAGAAATGACCTGAGGTGCGCGGTGAAGGGGCTGACTGTGACAGGTACACC	311
Qy	4580	AGGTCAAGAGAAAGATTCTGATGCCATCTTCAAGAAATGTGCTTGTCTCCACCGGCCCA	4639
Db	312	AGGCCAAGAGAAAGTGTGTAACCTGCAAGGCGCGTGCCTTACTCCACGGGGCCCA	371
Qy	4640	AAAGTGCAGATATGGATCTGAGATGGAGGACAAGGAGTGGGGCAAGATGATCTTTCAGG	4699
Db	372	AGGCGCGGCAATGGACCTGGAGTGGGGCCACAGGCCCGCATGCGCGCATCTCTGCAGG	431
Qy	4700	ATGAGACATCAACCAACCAAGTTGAGATGATTTGAAAGCACTGAACACACTGGGCCACT	4759
Db	432	ACGAGAGACGTACCAACCAAGTATGACACGATTGGAAAGAGGCTGAACACACTGGCTCACT	491
Qy	4760	ACCAAGTGCAGATGTCTCGTGGATGACATTAAGTCCAAACAGATGACAGCTATTAAC	4819
Db	432	ACCAAGTACACAGACGGGTCTCGGTGGACTGTGTGCCAAGACAGTCCGCTTACAAACA	551

QY 4820 CAGTGAACACTCCACCGCTCTCCAGACCTCAGCAAGTAATGAAAACTGATCCGGT 4879
DB 552 TCTCCAACTCTCCACCTTCAACCAAG--TCCCTCAGCAGATACGAGCAATGCTGCCA 608
QY 4880 AACAGGCGAGCCCGGACAGCTCCGCTCAGGACACCTATGATCACTCCTGACTGAGA 4939
DB 609 CGGCGACGAGCCCGGACAGCTCTGCTCCGCGACGCTTATGATACGCCCACTGAGGA 668
QY 4940 GTGAGTCAAGATGTGGACCTAGTAAGAACCCAGACGAGACCGAAGAGAGGAGG 4999
DB 669 GCGGACCAAGCTGTGGACCTGTGAAGAA-CACAGACCTGAGACGAGCTGAGGGTG 727
QY 5000 ACCGCGGAGCAGAAATGTGTCTGAAATCTACTTACCCGACTCCTGGCACTAGGGA 5059
DB 728 ACCGCGGAGCAGAAATGTGTCTGAAATCTACTTACCCGACTCCTGGCACTAGGGA 787
QY 5060 CACTGAGAAATTTGTGATGATCTCTTGAAGCACTTTGACAGCGCACCC-GTGGC 5118
DB 788 CACTGAGAAATTTGTGATGATCTCTTGAAGCACTTTGACAGCGCACCCGGGGGC 847
QY 5119 TCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5177
DB 848 TCGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 907
QY 5178 ACATGCAATTCATGACCCGACGCTCCGACCTGAGAGCAATTGCTGCTGCTGAG 5237
DB 908 GCACCAATTCATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 967
QY 5238 GTTTTGGG 5245
DB 968 CTTTCTGG 975

RESULT 15
BU149880 882 bp mRNA linear EST 03-SRP-2002
LOCUS BU149880
DEFINITION AGENCOURT 8095739 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6088019
5', mRNA sequence.
ACCESSION BU149880
VERSION BU149880.1 GI:22663412
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 882)
TITLE NIH-MGC http://mgs.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: gsa@ds-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LNCM325 row: 1 column: 12
High quality sequence stop: 631.
Location/Qualifiers
1..882
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6088019"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 112"
/note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directional cloning
strategy."</p></div>
<div data-bbox="742 522 815 943" data-label="Text">
<p>GGCAGAG(G). Library constructed by Ling Hong in the

Laboratory of Gerald M. Rubin (University of California,

Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and

SuperScript II RT (Life Technologies). Note: this is a

NH_MGC library."</p>
</div>
<div data-bbox="32 522 734 943" data-label="Text">
<p>ORIGIN

Query Match 10.2%; Score 578; DB 13; Length 882;

Best Local Similarity 78.5%; Pred. No. 5,6e-111;

Matches 692; Conservative 0; Mismatches 190; Indels 0; Gaps 0;</p>
<p>4363 GGGAGCCCTCTTCTCCCTGTTCTGTGCTGCTTCAAGCAGATGAGAAAGGCCCTCAT 4422

1 GGGAGCCCTCTTCTCCCTGTTCTGTGCTGCTTCAAGCAGATGAGAAAGGCCCTCAT 60

4423 GAGCCATCAAGGCGGAGGCGCGCTACTCCTTGAAGGAGGAGCAAGCTCATCCGACAG 4482

61 GATGCCATCAAGGCGGAGGCGCGCTACTCCTTGAAGGAGGAGCAAGCTCATCCGACAG 120

4483 ATTGACTTCAAAACCCCTGATCTGAGCTGTGAGCCGACAGATGCAAGTCCAGG 4542

121 ATGAGTACAAAGCCCTGATCTGAGCTGTGAGCCGACAGATGCAAGTCCAGG 180

4543 GTCCCATTAAGATCCCTCACTGTGACATGACATGATGCAAGGAGGAGTTCGAT 4602

181 ATCCAGTGAAGGTGTTAACTGTGACACATCAACAGGTCAAGGAGGAGTTCGAT 240

4603 GCCATCTTCAAGATGTGCTTGTCTCCACCGGCGCAAGCTGCAATATGATCTGAG 4662

241 GCCGTATTAAGATGTGCTTGTCTCCACCGGCGCAAGCTGCAATATGATCTGAG 300

4663 TGGGCAAGAGAGTGGGCGAGGATATCTTGAAGATGAACATCAACCAAGATT 4722

301 TGGGCGAAGGCGGATGCGCGGCGGCTGTGCTGCAATGAGATCAACCAAGATT 360

4723 GAGATGATTGAAGAGCAGTGAACACACTGAGCCCACTTCAAGGATGCGGATGCTCG 4782

361 GAGGTGATGAGAGCGGCTCAACACTGATGATGATGATGATGATGATGATGATGAT 420

4783 GTGCAATTAATGTCTCAAGCAGGTGACAGCTTAAAGCAGTGAACACTCAACGCTTC 4842

421 GTGCTCTGTGCTCCCAACAGACCTCTCTTCAACATCCCTGCTTGCAGATCTTC 480

4843 AGGACCTCAGCAAGTAAATGAAACATGATCCGATACAGGCGACCCCGACAGCTTC 4902

481 CGGAGCTCAGCAAGTAAATGAAACATGATCCGATACAGGCGACCCCGACAGCTTC 540

4903 CGCTCAGCAGACCTTATGATCTCTGACTGAGTGAAGTGAAGTGAAGTGAAGTGAAG 4962

541 CGGTCGCGGCGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 600

4963 GTGAAGAACAGCAGACAGCAGACAGCAGAGAGGAGGAGCGGAGGAGCAGAGATGCT 5022

601 GTGAAGAACAGCAGACAGCAGACAGCAGAGAGGAGGAGGAGGAGGAGGAGGAGG 660

5023 GAAATCTACCTGACCCGCTCTGAGCTTAAAGGAGCAGCTGCAAGAGTTTGTGATGAC 5082

661 GAGATCTACCTGACCCGCTCTGAGCTTAAAGGAGCAGCTGCAAGAGTTTGTGATGAC 720

5083 CTCTTTGAGCAATCTTCAAGACAGCAGACAGGAGGAGGAGGAGGAGGAGGAGGAG 5142

721 TTGTTTGAACCTTGTTCAGCACTGTGACCCGGGAGAGGCTCTTCCCTGCGCATGAG 780

5143 TACATGTTGACTTCTGTGATGACAGAGCTGATTAACATGAGCATTCATGACCCGACGTC 5202

781 TACATGTTGACTTCTGTGATGACAGAGCTGATTAACATGAGCATTCATGACAGATGTG 840

5203 CGCATTACTGAGAGAGCAATGCTGCTGCTGCTGAGGTTTGG 5244

841 CGGACACTGAGAAAGCACTGCTCCTTCTGCGCTTTCTGG 882</p>
</div>
<div data-bbox="0 522 19 765" data-label="Page-Footer">
<p>Search completed: February 19, 2004, 23:13:56</p>
</div>

Mon Feb 23 09:34:03 2004

us-09-964-956-12.rst

Page 28

Job time : 8756 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 23:18:35 / Search time 1415 Seconds

(without alignments)
17475.395 Million cell updates/sec

Title: US-09-964-956-12

Perfect score: 5691 1 atgaaagccatgccttgaaa.....gcttagacagctgaataaa 5691

Sequence: 1 atgaaagccatgccttgaaa.....gcttagacagctgaataaa 5691

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 2167151695 residues

Word size : 0
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

GenBml:.*
1: gb_ba:.*
2: gb_htg:.*
3: gb_in:.*
4: gb_cm:.*
5: gb_ov:.*
6: gb_pat:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_sts:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vl:.*
15: em_ba:.*
16: em_fun:.*
17: em_hum:.*
18: em_in:.*
19: em_mu:.*
20: em_om:.*
21: em_or:.*
22: em_ov:.*
23: em_pat:.*
24: em_ph:.*
25: em_pl:.*
26: em_ro:.*
27: em_sts:.*
28: em_un:.*
29: em_vl:.*
30: em_htg_hum:.*
31: em_htg_in:.*
32: em_htg_other:.*
33: em_htg_mus:.*
34: em_htg_pln:.*
35: em_htg_rtd:.*
36: em_htg_mam:.*
37: em_htg_vtl:.*
38: em_sy:.*
39: em_htgo_hum:.*
40: em_htgo_mus:.*
41: em_htgo_other:.*

Pred. No. is the number of results predicted by chance to have a

Result No.	Score	Query Match	Length	DB ID	Description
1	5691	100.0	5691	6	AX503782 Sequence
2	5532	97.2	6367	6	AX492982 Sequence
3	2237	39.3	5986	9	AB046770 Homo sapi
4	1371	24.1	2015	9	BC028744 Homo sapi
5	1371	24.1	3310	9	AK123428 Homo sapi
6	1188	20.9	179145	9	AC009785 Homo sapi
7	1108	19.5	3665	6	AX833678 Sequence
8	1108	19.5	3666	9	AK095606 Homo sapi
9	1106	19.4	188090	9	AC011625 Homo sapi
10	1068	18.8	179592	2	AC009723 Homo sapi
11	878	15.4	3556	6	HSMB805617
12	793	13.9	2597	6	AX574564 Sequence
13	793	13.9	2597	9	AY358850 Homo sapi
14	706	12.4	1029	9	HSMB802037
15	509	8.9	74655	2	AC025595 Homo sapi
16	491	8.6	601	6	AX088116 Sequence
17	479	8.4	3560	6	AX747045 Homo sapi
18	479	8.4	3560	9	AK091538 Homo sapi
19	272	4.8	198470	9	AC018643 Homo sapi
20	241	4.2	171557	2	AC147275
21	216	3.8	185272	9	AC105443
22	134	2.4	151153	9	AC009364 Homo sapi
23	123	2.4	183538	2	AC147272
24	123	2.2	171557	2	AC147276
25	74	1.3	74655	2	AC025595 Homo sapi
26	61	1.1	247656	2	AC112331
27	27	1.1	274459	2	AC093959 Rattus no
28	28	1.1	284400	2	AC096840 Rattus no
29	53	0.9	7022	10	AB073228 Mus muscu
30	30	0.9	172060	2	AC102171
31	49	0.9	6039	9	HSSEKXENB
32	48	0.8	144560	2	BX664737 Homo sapi
33	48	0.8	160656	2	AL592072 Homo sapi
34	48	0.8	219447	9	HUMFLNG6PD
35	47	0.8	4445	9	AK126101
36	47	0.8	5263	9	HSNOVPROT
37	47	0.8	5895	6	AX686475 Sequence
38	47	0.8	6147	6	AX704766 Homo sapi
39	47	0.8	6693	6	BD183298 Novel gen
40	47	0.8	91240	9	AC130566 Homo sapi
41	47	0.8	152135	9	AC011199 Homo sapi
42	47	0.8	221370	2	AC126205
43	47	0.8	22898	2	AC131554
44	45	0.8	465	11	HSCE7605
45	44	0.8	172060	2	AC102171
46	44	0.8	238662	2	AC102156 Mus muscu
47	39	0.7	1817	9	BC006193 Homo sapi
48	39	0.7	2875	9	BC009343 Homo sapi
49	39	0.7	3985	9	HSOCTPROT
50	39	0.7	6263	9	AB007932
51	39	0.7	6387	6	AX405895 Sequence
52	39	0.7	118429	9	AL158017 Homo sapi
53	39	0.7	198906	9	AL356275 Human DNA
54	38	0.7	776	4	AF176418 Bos tauru
55	38	0.7	6466	10	D86948 Mus musculu
56	38	0.7	211217	2	AC123387
57	35	0.6	218548	2	AC132672
58	35	0.6	228751	2	EX648286
59	35	0.6	229922	2	AX102141
60	35	0.6	230199	5	EX537358 Zebrafish
61	35	0.6	251269	2	AC145193 Gallus ga
62	35	0.6	260656	2	EX640582 Danio rer
63	34	0.6	5982	10	D86950 Mus musculu
64	34	0.6	188873	2	AC073658 Mus muscu
65	34	0.6	198653	2	AC091474 Mus muscu

66	34	0.6	200322	2	AL808145	Mus muscu	139	21	0.4	2815	10	BC011069
67	34	0.6	203189	10	AL807376	Mouse DNA	140	21	0.4	3078	10	WMPCTA12A
68	34	0.6	214505	2	AC025784	Mus muscu	141	21	0.4	3111	10	RN036444
69	34	0.6	220228	2	AC094668	Rattus no	142	21	0.4	5249	10	AC1122467
70	33	0.6	146405	2	AC122795	Mus muscu	143	21	0.4	46794	2	AC114683
71	33	0.6	255915	2	AC105876	Rattus no	144	21	0.4	60111	2	AC117617
72	32	0.6	1755	10	BC043322	Mus muscu	145	21	0.4	65513	2	AC103748
73	32	0.6	4320	10	AK128133	Mus muscu	146	21	0.4	66118	9	AL162259
74	32	0.6	8466	10	AB072381	Mus muscu	147	21	0.4	73947	9	AL445248
75	32	0.6	109148	2	AC138439	Taki-fugu	148	21	0.4	83559	9	AC044786
76	31	0.5	208593	2	AC140383	Mus muscu	149	21	0.4	98861	9	AL136961
77	31	0.5	229392	2	AC112372	Rattus no	150	21	0.4	99512	9	AC004606
78	31	0.5	238662	2	AC102156	Mus muscu	151	21	0.4	101458	2	AP000680
79	31	0.5	255661	2	AC121624	Rattus no	152	21	0.4	109602	2	AC131780
80	29	0.5	582	5	AB055678	Danio rer	153	21	0.4	109610	10	AC135241
81	29	0.5	768	6	AB055678	Sequence	154	21	0.4	109624	2	AC144935
82	29	0.5	768	6	BD145964	Primer fo	155	21	0.4	110000	2	AC109940
83	29	0.5	1012	6	BC032125	Homo sapi	156	21	0.4	115145	2	AL807240
84	29	0.5	1676	6	AB078694	Sequence	157	21	0.4	129414	2	AC143614
85	29	0.5	1767	6	BD157390	Primer fo	158	21	0.4	134506	9	HS357116
86	29	0.5	2294	9	AK094367	AK021813	159	21	0.4	135638	1	AF484556
87	29	0.5	2294	9	AK094367	AK094367	160	21	0.4	137908	10	AL607024
88	29	0.5	4196	9	AK128612	Homo sapi	161	21	0.4	142696	2	AC143051
89	26	0.5	26	6	AX503877	Sequence	162	21	0.4	147138	10	AL672094
90	26	0.5	6163	5	XELPLEX	D38175 Xenopus lae	163	21	0.4	151174	2	AC016189
91	26	0.5	75429	2	AC101416	Mus muscu	164	21	0.4	154896	2	AC110481
92	25	0.4	6330	6	BC056475	Sequence	165	21	0.4	158170	8	AC093180
93	25	0.4	6330	10	AB094949	Mus muscu	166	21	0.4	160989	8	OSJN00089
94	25	0.4	6330	10	AB094949	Mus muscu	167	21	0.4	165776	2	AC109157
95	25	0.4	6893	10	AK122289	Sequence	168	21	0.4	166804	8	OSJN00077
96	25	0.4	212744	2	AC106326	Rattus no	169	21	0.4	168639	2	AC121229
97	25	0.4	216192	10	AC111280	Rattus no	170	21	0.4	171368	9	HS095738
98	25	0.4	228202	5	BX284664	Zebrafish	171	21	0.4	172029	9	AC105416
99	24	0.4	1486	4	OCHEMOPEX	X16429 Oryzola	172	21	0.4	172506	2	AC105416
100	23	0.4	190	6	AX904782	Sequence	173	21	0.4	175745	2	BS530060
101	23	0.4	1758	9	BD040315	BD040315	174	21	0.4	175849	2	AL512366
102	23	0.4	147686	10	HSW802538	AL162013 Homo sapi	175	21	0.4	175965	2	AC131778
103	23	0.4	157250	2	AC027264	AC027264 Homo sapi	176	21	0.4	176329	2	AC118047
104	23	0.4	164057	9	AL590138	Human DNA	177	21	0.4	178525	9	AC018612
105	23	0.4	169305	9	AC021036	AC021036 Homo sapi	178	21	0.4	183133	2	AC024688
106	23	0.4	183353	10	AL935278	Mouse DNA	179	21	0.4	185973	2	AC128212
107	23	0.4	190748	9	AC010969	AC010969 Homo sapi	180	21	0.4	189268	9	AP000866
108	23	0.4	217392	2	AC107414	Rattus no	181	21	0.4	190104	2	AC122985
109	23	0.4	218448	2	BX324190	Danio rer	182	21	0.4	192729	2	AC147396
110	23	0.4	229132	2	AC112047	Rattus no	183	21	0.4	195961	10	AC124709
111	23	0.4	253343	2	AC102566	Sequence	184	21	0.4	198197	9	AC073957
112	22	0.4	6178	6	AX503876	AX503876	185	21	0.4	198197	10	AC124709
113	22	0.4	143406	10	AC092857	Rattus no	186	21	0.4	198197	10	AC124709
114	22	0.4	174938	10	AC121975	Mus muscu	187	21	0.4	198197	10	AC124709
115	22	0.4	182452	9	AL161771	Human DNA	188	21	0.4	198197	10	AC124709
116	22	0.4	190135	2	AC142039	Mus muscu	189	21	0.4	198197	10	AC124709
117	22	0.4	197333	2	AC120001	Mus muscu	190	21	0.4	198197	10	AC124709
118	22	0.4	202391	10	AC110040	Mus muscu	191	21	0.4	198197	10	AC124709
119	22	0.4	205587	2	AC139761	Mus muscu	192	21	0.4	198197	10	AC124709
120	22	0.4	207228	2	AC109215	Mus muscu	193	21	0.4	198197	10	AC124709
121	22	0.4	228678	2	AC110740	Mus muscu	194	21	0.4	198197	10	AC124709
122	22	0.4	230329	2	AC142070	Rattus no	195	21	0.4	198197	10	AC124709
123	22	0.4	232691	2	AC127871	Mus muscu	196	21	0.4	198197	10	AC124709
124	22	0.4	234882	2	AC127871	Mus muscu	197	21	0.4	198197	10	AC124709
125	22	0.4	237777	2	AC136165	Rattus no	198	21	0.4	198197	10	AC124709
126	22	0.4	244795	2	AC131632	Rattus no	199	21	0.4	198197	10	AC124709
127	22	0.4	246196	2	AC105641	Rattus no	200	21	0.4	198197	10	AC124709
128	22	0.4	254380	2	AC105841	Rattus no	201	21	0.4	198197	10	AC124709
129	22	0.4	273225	2	AC129440	Rattus no	202	21	0.4	198197	10	AC124709
130	22	0.4	273225	2	AC129440	Rattus no	203	21	0.4	198197	10	AC124709
131	22	0.4	349841	1	AX572606	Rhodospir	204	21	0.4	198197	10	AC124709
132	22	0.4	349841	1	AX572606	Rhodospir	205	21	0.4	198197	10	AC124709
133	22	0.4	349841	1	AX572606	Rhodospir	206	21	0.4	198197	10	AC124709
134	21	0.4	1874	10	AF540386	Mus muscu	207	21	0.4	198197	10	AC124709
135	21	0.4	1874	10	AF540386	Mus muscu	208	21	0.4	198197	10	AC124709
136	21	0.4	2591	10	RN036444	Rattus no	209	21	0.4	198197	10	AC124709
137	21	0.4	2591	10	RN036444	Rattus no	210	21	0.4	198197	10	AC124709
138	21	0.4	2669	10	BC062112	Mus muscu	211	21	0.4	198197	10	AC124709

C 212	21	0.4	244343	2	AC106516	Rattus no	C 285	20	0.4	2660	9	AK022651	AK022651 Homo sapi
C 213	21	0.4	245892	2	AC096884	Rattus no	C 286	20	0.4	2166	9	BC004542	BC004542 Homo sapi
C 214	21	0.4	247645	2	AC118405	Rattus no	C 287	20	0.4	2291	9	BC011892	BC011892 Homo sapi
C 215	21	0.4	250415	2	AC097418	Rattus no	C 288	20	0.4	2389	9	AP464935	AP464935 Homo sapi
C 216	21	0.4	250627	2	AC097031	Rattus no	C 289	20	0.4	2444	6	E36720	E36720 Novel trans
C 217	21	0.4	254288	2	AC126112	Rattus no	C 290	20	0.4	2522	6	AR339211	AR339211 Sequence
C 218	21	0.4	254708	2	AC111761	Rattus no	C 291	20	0.4	2533	10	BC004762	BC004762 Mus muscu
C 219	21	0.4	254708	2	AC115307	Rattus no	C 292	20	0.4	2564	6	AX176754	AX176754 Homo sapi
C 220	21	0.4	260600	2	AC121309	Rattus no	C 293	20	0.4	2564	6	AX091056	AX091056 Homo sapi
C 221	21	0.4	262886	2	AC121309	Rattus no	C 294	20	0.4	2775	10	AB041607	AB041607 Mus muscu
C 222	21	0.4	265567	2	AC096335	Rattus no	C 295	20	0.4	2935	10	BC003293	BC003293 Mus muscu
C 223	21	0.4	271040	2	AC121473	Rattus no	C 296	20	0.4	2971	6	BD127371	BD127371 Primer fo
C 224	21	0.4	276800	1	SC093115	Streptomy	C 297	20	0.4	2971	6	BD127371	BD127371 Primer fo
C 225	21	0.4	279593	2	AC114015	Rattus no	C 298	20	0.4	3038	9	AK126394	AK126394 Homo sapi
C 226	21	0.4	280025	2	AC108655	Rattus no	C 299	20	0.4	3038	9	AK126394	AK126394 Homo sapi
C 227	21	0.4	283782	2	AC096338	Rattus no	C 300	20	0.4	3184	10	BC051045	BC051045 Mus muscu
C 228	21	0.4	284396	2	AC105485	Rattus no	C 301	20	0.4	3504	9	AK056543	AK056543 Homo sapi
C 229	21	0.4	294396	2	AC105485	Rattus no	C 302	20	0.4	3726	9	AK025701	AK025701 Homo sapi
C 230	21	0.4	346542	2	AC120727	Rattus no	C 303	20	0.4	3849	6	AX058889	AX058889 Sequence
C 231	21	0.4	347368	2	AC096620	Rattus no	C 304	20	0.4	4125	8	CRE427966	CRE427966 Chlamydom
C 232	20	0.4	207	9	HS62A1R	H. sapiens	C 305	20	0.4	4355	10	BC054788	BC054788 Mus muscu
C 233	20	0.4	242	9	HS63MRNAG	H. sapiens m	C 306	20	0.4	4378	10	BC053092	BC053092 Mus muscu
C 234	20	0.4	272	9	HS179E11F	Z54966 H. sapiens C	C 307	20	0.4	4725	10	AF190578	AF190578 Homo sapi
C 235	20	0.4	311	6	AR251662	Sequence	C 308	20	0.4	5310	10	BC007481	BC007481 Homo sapi
C 236	20	0.4	318	6	BD004150	BD004150 5'SST and	C 309	20	0.4	6252	6	AX088110	AX088110 Sequence
C 237	20	0.4	324	6	AX008405	Sequence	C 310	20	0.4	6252	6	AX210612	AX210612 Sequence
C 238	20	0.4	328	6	AX260802	Sequence	C 311	20	0.4	6252	6	AX35107	AX35107 Sequence
C 239	20	0.4	332	10	AF029886	Rattus no	C 312	20	0.4	6252	6	AB002313	AB002313 Human mRN
C 240	20	0.4	381	9	HS4430739	AF029886 Rattus no	C 313	20	0.4	8960	1	AF173226	AF173226 Streptococ
C 241	20	0.4	382	6	AX895484	Sequence	C 314	20	0.4	8960	1	AF173226	AF173226 Streptococ
C 242	20	0.4	382	6	BD031017	Sequence	C 315	20	0.4	1158	1	AE004352	AE004352 Pseudomon
C 243	20	0.4	412	6	AX895483	Sequence	C 316	20	0.4	1158	1	AE004352	AE004352 Pseudomon
C 244	20	0.4	412	6	BD031016	Sequence	C 317	20	0.4	12744	1	AE004401	AE004401 Streptomy
C 245	20	0.4	419	6	AX396560	Sequence	C 318	20	0.4	1374	1	AE003997	AE003997 Xylella I
C 246	20	0.4	456	6	AX895482	Sequence	C 319	20	0.4	13842	6	BD217185	BD217185 DNA encod
C 247	20	0.4	456	6	BD031015	Sequence	C 320	20	0.4	18979	10	BX842567	BX842567 Mouse DNA
C 248	20	0.4	462	6	BD65166	Compounds	C 321	20	0.4	23518	9	AL589735	AL589735 Human DNA
C 249	20	0.4	462	6	AR401152	Sequence	C 322	20	0.4	29074	3	LMF05213T	LMF05213T Human DNA
C 250	20	0.4	462	6	AX192598	Sequence	C 323	20	0.4	32539	9	HS1LC2	HS1LC2 Human DNA
C 251	20	0.4	486	6	AX895485	Sequence	C 324	20	0.4	37948	1	AF079138	AF079138 Streptomy
C 252	20	0.4	486	6	BD031018	Sequence	C 325	20	0.4	37948	1	BD217173	BD217173 DNA encod
C 253	20	0.4	506	6	AX804284	Sequence	C 326	20	0.4	38506	6	BD232534	BD232534 Recombina
C 254	20	0.4	615	6	AX884867	Sequence	C 327	20	0.4	38506	6	AR271680	AR271680 Sequence
C 255	20	0.4	615	6	BD024477	Sequence	C 328	20	0.4	40548	8	AB011480	AB011480 Arabidops
C 256	20	0.4	616	6	BD229563	Human gen	C 329	20	0.4	4662	3	AC024755	AC024755 Genothad
C 257	20	0.4	635	8	AF362041	Dendrobilu	C 330	20	0.4	4858	9	AC016259	AC016259 Rattus no
C 258	20	0.4	670	6	AX899105	Sequence	C 331	20	0.4	52573	2	AC136259	AC136259 Rattus no
C 259	20	0.4	670	6	BD034638	Sequence	C 332	20	0.4	52573	2	AC074013	AC074013 Homo sapi
C 260	20	0.4	689	6	AX899102	Sequence	C 333	20	0.4	56557	9	AY115107	AY115107 Mus muscu
C 261	20	0.4	689	6	BD034635	Sequence	C 334	20	0.4	62297	10	AL359962	AL359962 Human DNA
C 262	20	0.4	700	6	AX410652	Sequence	C 335	20	0.4	67173	9	AP001505	AP001505 Homo sapi
C 263	20	0.4	700	6	BC006791	Sequence	C 336	20	0.4	70282	9	AP001505	AP001505 Homo sapi
C 264	20	0.4	700	9	HSU12404	BC006791 Homo sapi	C 337	20	0.4	70282	9	AP001505	AP001505 Homo sapi
C 265	20	0.4	709	9	BC011356	BC011356 Homo sapi	C 338	20	0.4	74194	9	AL356449	AL356449 Human DNA
C 266	20	0.4	711	6	BD189960	Tumor sht	C 339	20	0.4	74194	9	AL356449	AL356449 Human DNA
C 267	20	0.4	741	6	AB082926	Homo sapi	C 340	20	0.4	77588	8	AB016885	AB016885 Arabidops
C 268	20	0.4	744	6	AX899094	Sequence	C 341	20	0.4	79340	8	AP003956	AP003956 Oryza sat
C 269	20	0.4	744	6	BD034627	Sequence	C 342	20	0.4	79884	9	AC010944	AC010944 Homo sapi
C 270	20	0.4	750	6	BD220723	Human gen	C 343	20	0.4	85440	2	AC034264	AC034264 Mus muscu
C 271	20	0.4	834	6	AX8677417	Sequence	C 344	20	0.4	88421	6	AX417445	AX417445 Sequence
C 272	20	0.4	834	6	BD147479	Primer fo	C 345	20	0.4	88421	6	AX417445	AX417445 Sequence
C 273	20	0.4	1396	6	S76730	KM1-clone X	C 346	20	0.4	88421	6	AX417445	AX417445 Sequence
C 274	20	0.4	1403	10	BC024509	BC024509 Mus muscu	C 347	20	0.4	95599	2	AC110929	AC110929 (8 of
C 275	20	0.4	1413	9	BT006887	BT006887 Homo sapi	C 348	20	0.4	95599	2	AC110929	AC110929 (8 of
C 276	20	0.4	1413	12	BT007594	BT007594 Homo sapi	C 349	20	0.4	95599	2	AC110929	AC110929 (8 of
C 277	20	0.4	1618	10	MM058887	US8887 Mus musculu	C 350	20	0.4	95599	2	AC110929	AC110929 (8 of
C 278	20	0.4	1636	6	AR175265	AR175265 Sequence	C 351	20	0.4	95599	2	AC110929	AC110929 (8 of
C 279	20	0.4	1636	10	BC027096	BC027096 Mus muscu	C 352	20	0.4	95599	2	AC110929	AC110929 (8 of
C 280	20	0.4	1782	10	AF077738	AF077738 Rat nortve	C 353	20	0.4	95599	2	AC110929	AC110929 (8 of
C 281	20	0.4	1850	9	BC061993	BC061993 Rattus no	C 354	20	0.4	95599	2	AC110929	AC110929 (8 of
C 282	20	0.4	1876	10	HSMB000181	AL050280 Homo sapi	C 355	20	0.4	95599	2	AC110929	AC110929 (8 of
C 283	20	0.4	1946	6	AK025415	AK025415 Homo sapi	C 356	20	0.4	95599	2	AC110929	AC110929 (8 of
C 284	20	0.4	2060	6	AX879531	AX879531 Sequence	C 357	20	0.4	95599	2	AC110929	AC110929 (8 of
C 284	20	0.4	2060	6	BD157886	BD157886 Primer fo	C 357	20	0.4	95599	2	AC110929	AC110929 (8 of

358	C	358	0.4	110000	2	LMFLCR36_23	Continuation (24 o	C 431	20	0.4	179591	2	AC112970	AC112970 Mus muscu
359	C	359	0.4	112339	2	HS110F11	AL033526 Human DNA	C 432	20	0.4	179756	2	AC074039	AC074039 Mus muscu
360	C	360	0.4	112714	2	AC144689	AC144689 Ornithoth	C 433	20	0.4	180239	10	AL671190	AL671190 Mouse DNA
361	C	361	0.4	112846	2	AC004991	AC004991 Homo sapi	C 434	20	0.4	180690	10	AC121922	AC121922 Mus muscu
362	C	362	0.4	114774	2	AL603748	AL603748 Homo sapi	C 435	20	0.4	181388	2	AC130062	AC130062 Rat
363	C	363	0.4	115292	9	AC108742	AC108742 Homo sapi	C 436	20	0.4	181808	2	AC362627	AC362627 Homo sapi
364	C	364	0.4	115295	9	AL136967	AL136967 Human DNA	C 437	20	0.4	183440	2	AC069295	AC069295 Homo sapi
365	C	365	0.4	117000	9	AC117475	AC117475 Homo sapi	C 438	20	0.4	183462	10	AC126675	AC126675 Mus muscu
366	C	366	0.4	118299	10	AL929069	AL929069 Mouse DNA	C 439	20	0.4	184467	2	AC073241	AC073241 Homo sapi
367	C	367	0.4	120323	2	BX539322	BX539322 Dantio rer	C 440	20	0.4	184590	9	AL390059	AL390059 Human DNA
368	C	368	0.4	126982	2	BX649443	BX649443 Homo sapi	C 441	20	0.4	184686	2	AC136693	AC136693 Gallus ga
369	C	369	0.4	128720	2	AC124970	AC124970 Medicago	C 442	20	0.4	184758	2	AC122546	AC122546 Mus muscu
370	C	370	0.4	129519	10	AL645951	AL645951 Mouse DNA	C 443	20	0.4	185399	2	AC026840	AC026840 Homo sapi
371	C	371	0.4	130110	9	AL731777	AL731777 Human DNA	C 444	20	0.4	185399	2	AC101747	AC101747 Mus muscu
372	C	372	0.4	133915	9	AC108740	AC108740 Homo sapi	C 445	20	0.4	185788	2	AC136510	AC136510 Pan trogl
373	C	373	0.4	134666	2	AC140091	AC140091 Bos tauru	C 446	20	0.4	185879	2	AC135705	AC135705 Rat
374	C	374	0.4	135692	9	AC021134	AC021134 Homo sapi	C 447	20	0.4	186285	2	AC140098	AC140098 Sus scrofa
375	C	375	0.4	136687	10	AC005816	AC005816 Mus muscu	C 448	20	0.4	186795	2	AC090492	AC090492 Mus muscu
376	C	376	0.4	137506	9	AC004824	AC004824 Homo sapi	C 449	20	0.4	187117	9	AC016749	AC016749 Homo sapi
377	C	377	0.4	140113	2	AC116824	AC116824 Mus muscu	C 450	20	0.4	188417	2	AC025205	AC025205 Mus muscu
378	C	378	0.4	140630	2	AL357514	AL357514 Human DNA	C 451	20	0.4	188873	2	AC073658	AC073658 Mus muscu
379	C	379	0.4	140894	2	AC144754	AC144754 Ornithoth	C 452	20	0.4	189118	2	AC133835	AC133835 Mus muscu
380	C	380	0.4	141854	2	AC021595	AC021595 Homo sapi	C 453	20	0.4	189846	2	AC140443	AC140443 Mus muscu
381	C	381	0.4	143727	2	AP005723	AP005723 Oryza sat	C 454	20	0.4	189939	2	AC109273	AC109273 Mus muscu
382	C	382	0.4	146651	2	AC060778	AC060778 Homo sapi	C 455	20	0.4	190309	2	AC147151	AC147151 Pan trogl
383	C	383	0.4	150240	2	AC146458	AC146458 Pan trogl	C 456	20	0.4	190997	2	BX649409	BX649409 Dantio rer
384	C	384	0.4	150724	10	AP003150	AP003150 Mus muscu	C 457	20	0.4	191503	2	AC110561	AC110561 Homo sapi
385	C	385	0.4	152337	8	AY325816	AY325816 Zee maye	C 458	20	0.4	191701	2	AC080189	AC080189 Homo sapi
386	C	386	0.4	152395	9	AC123022	AC123022 Homo sapi	C 459	20	0.4	193189	2	AC113069	AC113069 Mus muscu
387	C	387	0.4	152449	2	AC080175	AC080175 Homo sapi	C 460	20	0.4	193311	2	AC134403	AC134403 Mus muscu
388	C	388	0.4	152992	2	AC023804	AC023804 Mus muscu	C 461	20	0.4	193577	8	AP004365	AP004365 Oryza sat
389	C	389	0.4	153736	2	AC141187	AC141187 Rattus no	C 462	20	0.4	193709	2	AC074205	AC074205 Mus muscu
390	C	390	0.4	154412	2	HS144F13	HS144F13 Human DNA	C 463	20	0.4	194305	2	AC117694	AC117694 Mus muscu
391	C	391	0.4	154741	10	AC124694	AC124694 Mus muscu	C 464	20	0.4	195513	2	AC133328	AC133328 Rattus no
392	C	392	0.4	155349	2	AL162257	AL162257 Homo sapi	C 465	20	0.4	196146	2	AC023609	AC023609 Mus muscu
393	C	393	0.4	155818	9	AL954253	AL954253 Pan trogl	C 466	20	0.4	197201	2	AC119091	AC119091 Homo sapi
394	C	394	0.4	156124	10	AL125325	AL125325 Mus muscu	C 467	20	0.4	199103	2	AC080076	AC080076 Homo sapi
395	C	395	0.4	156393	2	AC116177	AC116177 Homo sapi	C 468	20	0.4	200348	9	AC005516	AC005516 Homo sapi
396	C	396	0.4	157088	10	AC127292	AC127292 Mus muscu	C 469	20	0.4	200512	2	AC135866	AC135866 Rattus no
397	C	397	0.4	157168	10	AL772232	AL772232 Mus muscu	C 470	20	0.4	201339	2	AC135866	AC135866 Rattus no
398	C	398	0.4	157790	9	AL133413	AL133413 Human DNA	C 471	20	0.4	203071	2	AC120394	AC120394 Mus muscu
399	C	399	0.4	158037	2	AC010131	AC010131 Homo sapi	C 472	20	0.4	203106	10	AC021667	AC021667 Mus muscu
400	C	400	0.4	158573	2	AC078823	AC078823 Homo sapi	C 473	20	0.4	203560	2	AC026298	AC026298 Homo sapi
401	C	401	0.4	159499	9	AC146456	AC146456 Pan trogl	C 474	20	0.4	205053	9	AC011749	AC011749 Homo sapi
402	C	402	0.4	160178	9	AL162584	AL162584 Human DNA	C 475	20	0.4	205819	2	AC106805	AC106805 Homo sapi
403	C	403	0.4	162701	2	AC073317	AC073317 Homo sapi	C 476	20	0.4	206470	2	AC135811	AC135811 Rattus no
404	C	404	0.4	162956	2	AC068168	AC068168 Homo sapi	C 477	20	0.4	206592	10	AL122471	AL122471 Mus muscu
405	C	405	0.4	162936	2	AC051655	AC051655 Homo sapi	C 478	20	0.4	207265	2	AL669950	AL669950 Mouse DNA
406	C	406	0.4	162932	2	AC025624	AC025624 Homo sapi	C 479	20	0.4	207265	2	AC122327	AC122327 Mus muscu
407	C	407	0.4	164632	2	AC140690	AC140690 Rattus no	C 480	20	0.4	208207	10	AC145450	AC145450 Mus muscu
408	C	408	0.4	165353	2	BX649446	BX649446 Dantio rer	C 481	20	0.4	208404	10	AC103564	AC103564 Mus muscu
409	C	409	0.4	165890	2	AC119710	AC119710 Rattus no	C 482	20	0.4	209247	2	AC106113	AC106113 Homo sapi
410	C	410	0.4	167710	2	AC109891	AC109891 Rattus no	C 483	20	0.4	210844	2	AC129214	AC129214 Mus muscu
411	C	411	0.4	167925	2	AC142899	AC142899 Macaca mu	C 484	20	0.4	212844	2	AC078978	AC078978 Homo sapi
412	C	412	0.4	167962	2	AC073484	AC073484 Homo sapi	C 485	20	0.4	212844	2	AC117723	AC117723 Mus muscu
413	C	413	0.4	168813	9	AP000555	AP000555 Homo sapi	C 486	20	0.4	213056	2	AC117723	AC117723 Mus muscu
414	C	414	0.4	168981	10	AL591469	AL591469 Mouse DNA	C 487	20	0.4	214042	10	AC127556	AC127556 Mus muscu
415	C	415	0.4	169649	2	AC046163	AC046163 Homo sapi	C 488	20	0.4	215795	2	BX649437	BX649437 Dantio rer
416	C	416	0.4	169741	9	AC097468	AC097468 Homo sapi	C 489	20	0.4	215826	10	AC129078	AC129078 Mus muscu
417	C	417	0.4	170245	9	HS109F14	HS109F14 Human DNA	C 490	20	0.4	215895	2	AC097998	AC097998 Rattus no
418	C	418	0.4	170525	2	AC101804	AC101804 Mus muscu	C 491	20	0.4	215967	2	AC120425	AC120425 Mus muscu
419	C	419	0.4	170652	2	AC092934	AC092934 Homo sapi	C 492	20	0.4	216314	2	AC110164	AC110164 Mus muscu
420	C	420	0.4	171052	2	AC025802	AC025802 Homo sapi	C 493	20	0.4	216966	2	AC111973	AC111973 Rattus no
421	C	421	0.4	171364	8	OSJN00025	OSJN00025 Oryza sat	C 494	20	0.4	218018	2	AC123866	AC123866 Mus muscu
422	C	422	0.4	173389	9	CNS05TCM	AL605694 Oryza sat	C 495	20	0.4	218242	2	AC105860	AC105860 Rattus no
423	C	423	0.4	173700	9	AC096919	AC096919 Homo sapi	C 496	20	0.4	219036	2	AC105588	AC105588 Homo sapi
424	C	424	0.4	173837	10	AC124681	AC124681 Mus muscu	C 497	20	0.4	219443	2	AC023330	AC023330 Rattus no
425	C	425	0.4	173956	9	AC105383	AC105383 Homo sapi	C 498	20	0.4	219649	2	AC133222	AC133222 Rattus no
426	C	426	0.4	175066	9	AC005900	AC005900 Homo sapi	C 499	20	0.4	220524	2	AC103065	AC103065 Rattus no
427	C	427	0.4	176027	2	AL732575	AL732575 Mus muscu	C 500	20	0.4	220887	10	AC124741	AC124741 Mus muscu
428	C	428	0.4	176144	2	AC012408	AC012408 Homo sapi	C 501	20	0.4	220930	2	AC135137	AC135137 Rattus no
429	C	429	0.4	176629	9	AC008641	AC008641 Homo sapi	C 502	20	0.4	221753	2	AC114835	AC114835 Rattus no
430	C	430	0.4	179276	2	AC067884	AC067884 Homo sapi	C 503	20	0.4	221859	10	AL732309	AL732309 Mouse DNA

C 504	20	0.4 222182	2	AC123412	Rattus no	C 577	20	0.4 275900	2	AC123150	AC123150 Rattus no
505	20	0.4 222589	10	AC113364	Mus muscu	C 578	20	0.4 281085	2	AC123455	AC123455 Rattus no
506	20	0.4 222639	2	AC130083	Rattus no	C 579	20	0.4 284347	2	AC135290	AC135290 Mus muscu
507	20	0.4 223911	2	AC136740	Mus muscu	C 580	20	0.4 285152	2	AC116112	AC116112 Mus muscu
508	20	0.4 224138	10	AC133573	Mus muscu	C 581	20	0.4 286451	2	AC115130	AC115130 Rattus no
509	20	0.4 224764	2	AC122263	Mus muscu	C 582	20	0.4 289103	2	AC126288	AC126288 Rattus no
510	20	0.4 224859	2	AC073800	Mus muscu	C 583	20	0.4 290753	2	AC123256	AC123256 Rattus no
511	20	0.4 225012	2	AC113462	Mus muscu	C 584	20	0.4 292100	1	SC0939121	AC123126 Rattus no
512	20	0.4 225403	2	AC140451	Mus muscu	C 585	20	0.4 300029	1	AE016842	AE016842 Salomoni
513	20	0.4 225509	2	AC073782	Mus muscu	C 586	20	0.4 303688	10	AE016754	AE016754 Mus muscu
514	20	0.4 225541	2	AC126700	Rattus no	C 587	20	0.4 312626	2	AC109713	AC109713 Rattus no
515	20	0.4 225782	2	AC109529	Rattus no	C 588	20	0.4 325265	2	AC118802	AC118802 Rattus no
516	20	0.4 225788	2	AC123606	Mus muscu	C 589	20	0.4 325265	2	AC118802	AC118802 Rattus no
517	20	0.4 228051	2	AC097434	Rattus no	C 590	20	0.4 340000	9	HS21C101	HS21C101 Mus muscu
518	20	0.4 228178	2	AC111655	Rattus no	C 591	20	0.4 348764	1	BX569689	BX569689 Synechoco
519	20	0.4 228413	2	AC132453	Mus muscu	C 592	19	0.3 382	11	BT129486	BT129486 Bos taurus
520	20	0.4 228852	2	AC131526	Rattus no	C 593	19	0.3 382	11	AY254001	AY254001 Sus scrof
521	20	0.4 229909	2	AC111310	Rattus no	C 594	19	0.3 426	9	HS4240590	AJ240590 Homo sapi
522	20	0.4 230188	2	AC091604	Mus muscu	C 595	19	0.3 429	10	FE257137S07	FE257137 Mus muscu
523	20	0.4 233180	2	AC134135	Rattus no	C 596	19	0.3 503	5	AB062427	AB062427 Anguilla
524	20	0.4 233510	2	AC136569	Rattus no	C 597	19	0.3 503	6	BD012143	BD012143 Novel pep
525	20	0.4 233718	2	AC094024	Rattus no	C 598	19	0.3 529	11	G58051	G58051 SHGC-103874
526	20	0.4 234053	2	AC128262	Rattus no	C 599	19	0.3 539	11	G97698	G97698 S208P6575FE
527	20	0.4 234300	2	AC118431	Rattus no	C 600	19	0.3 545	4	AF381310	AF381310 Ornithoth
528	20	0.4 235050	2	AC096077	Rattus no	C 601	19	0.3 553	1	AF448095	AF448095 Synechoco
529	20	0.4 235427	2	AC140082	Rattus no	C 602	19	0.3 560	1	AF245157	AF245157 Synechoco
530	20	0.4 236159	2	AC130109	Rattus no	C 603	19	0.3 560	1	AF245158	AF245158 Synechoco
531	20	0.4 237304	2	AC096114	Rattus no	C 604	19	0.3 561	1	AF448096	AF448096 Synechoco
532	20	0.4 238007	10	AC096114	Rattus no	C 605	19	0.3 565	1	AF245125	AF245125 Synechoco
533	20	0.4 239628	2	AC128574	Rattus no	C 606	19	0.3 567	1	AF245154	AF245154 Synechoco
534	20	0.4 240203	2	AC128562	Rattus no	C 607	19	0.3 570	11	HSC01A04	AL158551 H.sapiens
535	20	0.4 240362	2	AC132653	Mus muscu	C 608	19	0.3 572	1	AF245155	AF245155 Synechoco
536	20	0.4 241781	2	AC122087	Rattus no	C 609	19	0.3 588	6	AR093091	AR093091 Sequence
537	20	0.4 242000	2	AC127751	Rattus no	C 610	19	0.3 588	6	BD009915	BD009915 BH3 inter
538	20	0.4 242482	2	AC115149	Rattus no	C 611	19	0.3 612	1	SSU52342-	U52343 Synechococ
539	20	0.4 242531	10	AC108805	Mus muscu	C 612	19	0.3 612	1	SSU52343	U52343 Synechococ
540	20	0.4 242662	2	AC080020	Mus muscu	C 613	19	0.3 648	14	AF3677124	AF3677124 Foot-and-
541	20	0.4 243047	2	AC126579	Rattus no	C 614	19	0.3 658	9	AY449016	AY449016 Autos aza
542	20	0.4 243269	2	AC117152	Rattus no	C 615	19	0.3 658	9	AY449015	AY449015 Ateles fu
543	20	0.4 243378	10	AC1222324	Mus muscu	C 616	19	0.3 658	9	BD125041	BD125041 Primer fo
544	20	0.4 243770	2	AC120763	Rattus no	C 617	19	0.3 739	6	BD126697	BD126697 Primer fo
545	20	0.4 244814	2	AC098223	Rattus no	C 618	19	0.3 739	11	BV044442	BV044442 S212P6045
546	20	0.4 245790	2	AC130142	Rattus no	C 619	19	0.3 759	6	BD126697	BD126697 Primer fo
547	20	0.4 246238	2	AC120998	Rattus no	C 620	19	0.3 779	11	BV063339	BV063339 S212P6040
548	20	0.4 246921	2	AC107248	Rattus no	C 621	19	0.3 791	6	AR2711766	AR2711766 Sequence
549	20	0.4 247108	2	AC123333	Rattus no	C 622	19	0.3 791	10	MM075506	MM075506 Mus muscu
550	20	0.4 248417	2	AC130165	Rattus no	C 623	19	0.3 796	9	HS4340034	AJ340034 Homo sapi
551	20	0.4 248938	2	AC095006	Rattus no	C 624	19	0.3 867	1	AB014978	AB014978 Pseudomon
552	20	0.4 249245	2	AC110695	Rattus no	C 625	19	0.3 900	6	141261	141261 Sequence 1
553	20	0.4 250335	2	AC114441	Rattus no	C 626	19	0.3 900	6	161439	161439 Sequence 1
554	20	0.4 251813	2	AC123454	Rattus no	C 627	19	0.3 900	6	172547	172547 Sequence 1
555	20	0.4 252110	2	AC094353	Rattus no	C 628	19	0.3 900	6	AB056724	AB056724 Rattus no
556	20	0.4 252404	2	AC097117	Rattus no	C 629	19	0.3 1157	8	EIN18591	AJ18591 Eleusine
557	20	0.4 252411	2	AC094615	Rattus no	C 630	19	0.3 1170	8	CD4379036	AJ279036 Cynodon d
558	20	0.4 252839	2	AC095096	Rattus no	C 631	19	0.3 1188	8	ETB318590	AJ18590 Eragrosti
559	20	0.4 253874	2	AC131560	Rattus no	C 632	19	0.3 1188	8	WRB318585	AJ318585 Melinis t
560	20	0.4 254483	2	AC095874	Rattus no	C 633	19	0.3 1188	8	FGB318584	AJ318584 Paspali d
561	20	0.4 257210	2	AC098622	Rattus no	C 634	19	0.3 1188	8	FWA318586	AJ318586 Panichum m
562	20	0.4 257398	2	AC131885	Rattus no	C 635	19	0.3 1236	6	AR320776	AR320776 Sequence
563	20	0.4 259807	2	AC131559	Rattus no	C 636	19	0.3 1236	6	AX489396	AX489396 Sequence
564	20	0.4 261065	2	AC128913	Rattus no	C 637	19	0.3 1340	8	AK121530	AK121530 Oryza sat
565	20	0.4 261224	10	AL645527	Mouse DNA	C 638	19	0.3 1370	3	AK115471	AK115471 Clona int
566	20	0.4 263239	2	AC121737	Rattus no	C 639	19	0.3 1443	6	AF182168	AF182168 Rattus no
567	20	0.4 263297	2	AC134030	Rattus no	C 640	19	0.3 1467	6	AX698013	AX698013 Sequence
568	20	0.4 263776	2	AC087335	Mus muscu	C 641	19	0.3 1477	1	EIN3981	AJ03981 Enterobac
569	20	0.4 265136	2	AC094847	Rattus no	C 642	19	0.3 1515	4	SSIGFBP112	U21118 Sus scrofa
570	20	0.4 267462	2	AC131806	Rattus no	C 643	19	0.3 1562	10	BC025092	BC025092 Mus muscu
571	20	0.4 268050	1	AL637286	Salomoni	C 644	19	0.3 1617	10	BC018162	BC018162 Mus muscu
572	20	0.4 270105	1	AC099650	Homo sapi	C 645	19	0.3 1689	5	BC034240	BC034240 Mus muscu
573	20	0.4 270668	2	AC096354	Rattus no	C 646	19	0.3 1749	5	AB045224	AB045224 Gallus ga
574	20	0.4 272677	2	AC097289	Rattus no	C 647	19	0.3 1763	10	MM0277454	AJ277454 Mus Muscu
575	20	0.4 274703	2	AC110414	Rattus no	C 648	19	0.3 1886	9	AF131827	AF131827 Homo sapi
576	20	0.4 275507	2	AC099381	Rattus no	C 649	19	0.3 1906	10	BC002031	BC002031 Mus muscu

C 650	19	0.3	1955	9	HSU04209	U04209 Human aassoc	C 723	19	0.3	14735	10	AF317901	AF317901 Mus muscu
C 651	19	0.3	1963	3	SCNAAL1	X55439 S. gregaria	C 724	19	0.3	17042	1	AE006932	AE006932 Mycobacte
C 652	19	0.3	1966	9	BC023557	BC023557 Homo sapi	C 725	19	0.3	18510	9	U72787	U72787 Homo sapien
C 653	19	0.3	1992	4	RABTCBCAI	M4577 Rabbitt T-ce	C 726	19	0.3	20300	6	AX45137	AX45137 Sequence
C 654	19	0.3	2000	6	AX656080	AX656080 Sequence	C 727	19	0.3	22996	9	AC087175	AC087175 Homo sapi
C 655	19	0.3	2002	10	BC033315	BC033315 Mus muscu	C 728	19	0.3	26919	9	AP000335	AP000335 Homo sapi
C 656	19	0.3	2052	9	BC050742	BC050742 Homo sapi	C 729	19	0.3	27555	2	AC012903	AC012903 Drosophi
C 657	19	0.3	2072	9	AF145606	AF145606 Drosophi	C 730	19	0.3	28235	2	AC014089	AC014089 Drosophi
C 658	19	0.3	2170	10	AB100266	AB100266 Mus muscu	C 731	19	0.3	30310	6	AR271852	AR271852 Sequence
C 659	19	0.3	2238	4	RABNANUCCOT	MA4020 Oryctolagus	C 732	19	0.3	34010	1	R5PNER234	R5PNER234 Sequence
C 660	19	0.3	2238	6	111726	111726 Sequence 1	C 733	19	0.3	38192	1	BX530088	BX530088 Human DNA
C 661	19	0.3	2304	10	RAT3AHPD	M64393 Rat 3-alpha	C 734	19	0.3	39730	1	MTCY8D5	MTCY8D5 Human DNA
C 662	19	0.3	2345	10	AF077765	AF077765 Mus muscu	C 735	19	0.3	40806	9	MSGY140	MSGY140 Mycobacte
C 663	19	0.3	2350	9	HUMPTLL1	M4348 Human parat	C 736	19	0.3	43385	2	BX537148	BX537148 Human DNA
C 664	19	0.3	2404	9	AK025131	AK025131 Homo sapi	C 737	19	0.3	43557	2	AC147080	AC147080 Homo sapi
C 665	19	0.3	2414	6	BD127385	BD127385 Primer fo	C 738	19	0.3	43567	2	AC102848	AC102848 Homo sapi
C 666	19	0.3	2414	9	AK074960	AK074960 Homo sapi	C 739	19	0.3	45276	9	ALU12654	ALU12654 Human DNA
C 667	19	0.3	2444	6	AX713537	AX713537 Sequence	C 740	19	0.3	46651	9	AC000404	AC000404 Genomic s
C 668	19	0.3	2444	9	AK055075	AK055075 Homo sapi	C 741	19	0.3	47782	2	AC068445	AC068445 Homo sapi
C 669	19	0.3	2466	10	RAT53AD	D17310 Rattus norv	C 742	19	0.3	49850	2	AC139554	AC139554 Homo sapi
C 670	19	0.3	2477	6	AR410736	AR410736 Sequence	C 743	19	0.3	51705	6	AX695614	AX695614 Sequence
C 671	19	0.3	2477	6	AX464198	AX464198 Sequence	C 744	19	0.3	53242	9	AL365494	AL365494 Human DNA
C 672	19	0.3	2477	6	AX697578	AX697578 Sequence	C 745	19	0.3	55818	9	AC112228	AC112228 Homo sapi
C 673	19	0.3	2477	6	BD075807	BD075807 Secretary	C 746	19	0.3	58198	9	AC073524	AC073524 Homo sapi
C 674	19	0.3	2477	6	BD172367	BD172367 Secretary	C 747	19	0.3	60190	8	AP004786	AP004786 Oryza sat
C 675	19	0.3	2477	6	BD172686	BD172686 Secretd	C 748	19	0.3	60196	6	AX697977	AX697977 Sequence
C 676	19	0.3	2477	6	BD173005	BD173005 Secretd	C 749	19	0.3	62370	2	AL451008	AL451008 Homo sapi
C 677	19	0.3	2477	6	BD173324	BD173324 Secretd	C 750	19	0.3	62370	2	AL360009	AL360009 Homo sapi
C 678	19	0.3	2477	6	BD173358	BD173358 Secretary	C 751	19	0.3	62380	10	AL928702	AL928702 Mouse DNA
C 679	19	0.3	2477	6	AX358496	AX358496 Homo sapi	C 752	19	0.3	63229	2	AC111184	AC111184 Homo sapi
C 680	19	0.3	2530	9	AX127587	AX127587 Homo sapi	C 753	19	0.3	63525	2	AC108463	AC108463 Homo sapi
C 681	19	0.3	2764	6	AX1713968	AX1713968 Sequence	C 754	19	0.3	63598	2	OS1G00045	OS1G00045 Oryza sat
C 682	19	0.3	2886	8	AX251482	AX251482 Echinochl	C 755	19	0.3	64513	2	AC101466	AC101466 Mus muscu
C 683	19	0.3	2886	8	AX833475	AX833475 Sequence	C 756	19	0.3	64513	2	AC101466	AC101466 Mus muscu
C 684	19	0.3	2942	6	AK095247	AK095247 Homo sapi	C 757	19	0.3	64824	2	AC067782	AC067782 Homo sapi
C 685	19	0.3	2942	6	AK095247	AK095247 Homo sapi	C 758	19	0.3	65796	2	AC124642	AC124642 Mus muscu
C 686	19	0.3	2997	6	AX113389	AX113389 Sequence	C 759	19	0.3	67211	2	AC133284	AC133284 Mus muscu
C 687	19	0.3	2997	6	AK054713	AK054713 Homo sapi	C 760	19	0.3	67211	2	AC133284	AC133284 Mus muscu
C 688	19	0.3	3458	6	AX1714243	AX1714243 Sequence	C 761	19	0.3	67718	2	AC135349	AC135349 Homo sapi
C 689	19	0.3	3458	6	AK056454	AK056454 Homo sapi	C 762	19	0.3	68010	2	AC101197	AC101197 Mus muscu
C 690	19	0.3	3671	6	AX074078	AX074078 Sequence	C 763	19	0.3	68052	2	AC100691	AC100691 Homo sapi
C 691	19	0.3	3734	9	HSMB07851	AB047705 Homo sapi	C 764	19	0.3	68417	2	AC040891	AC040891 Homo sapi
C 692	19	0.3	3761	9	AB094095	AB094095 Homo sapi	C 765	19	0.3	69333	2	AC109778	AC109778 Homo sapi
C 693	19	0.3	3844	10	MMNADPTRH	Z49204 M. musculus	C 766	19	0.3	69586	2	AC115076	AC115076 Mus muscu
C 694	19	0.3	4121	10	AK123006	AK123006 Mus muscu	C 767	19	0.3	70089	2	AC027491	AC027491 Homo sapi
C 695	19	0.3	4142	10	BC008518	BC008518 Mus muscu	C 768	19	0.3	70574	2	AC136312	AC136312 Homo sapi
C 696	19	0.3	4474	6	158526	158526 Sequence 9	C 769	19	0.3	71517	9	BS000067	BS000067 Pan trogl
C 697	19	0.3	4474	6	187001	187001 Sequence 9	C 770	19	0.3	74098	2	AC1073862	AC1073862 Homo sapi
C 698	19	0.3	4474	6	AX036110	AX036110 Sequence	C 771	19	0.3	74530	2	AC108315	AC108315 Pan trogl
C 699	19	0.3	4784	6	AR248295	AR248295 Mus muscu	C 772	19	0.3	74687	10	AL683882	AL683882 Mouse DNA
C 700	19	0.3	4809	5	AP254792	AP254792 Dario rex	C 773	19	0.3	77155	9	AC117478	AC117478 Homo sapi
C 701	19	0.3	5050	9	HMWD43	M61827 Human leuko	C 774	19	0.3	80627	2	AC027381	AC027381 Mus muscu
C 702	19	0.3	5295	10	BC060277	BC060277 Mus muscu	C 775	19	0.3	80656	2	AC110616	AC110616 Homo sapi
C 703	19	0.3	5528	10	AK123377	AK123377 Mus muscu	C 776	19	0.3	81105	2	AC109086	AC109086 Mus muscu
C 704	19	0.3	5693	10	BC036727	BC036727 Mus muscu	C 777	19	0.3	82179	2	AC100650	AC100650 Homo sapi
C 705	19	0.3	5939	10	AB011534	AB011534 Rattus no	C 778	19	0.3	82414	2	AC129265	AC129265 Rattus no
C 706	19	0.3	5939	9	HSMB06877	BSX40787 Homo sapi	C 779	19	0.3	83433	2	AC137046	AC137046 Homo sapi
C 707	19	0.3	6503	9	HSSTAL	X52075 Human gene	C 780	19	0.3	83989	9	AL603713	AL603713 Human DNA
C 708	19	0.3	6647	9	HSMB08937	BX449786 Homo sapi	C 781	19	0.3	84364	8	HS52237	HS52237 Human DNA
C 709	19	0.3	7161	9	AB011168	AB011168 Homo sapi	C 782	19	0.3	85795	8	AP006090	AP006090 Lotus cor
C 710	19	0.3	7174	9	HSMB04580	AB011367 Homo sapi	C 783	19	0.3	86451	2	AC107037	AC107037 Homo sapi
C 711	19	0.3	8388	9	HS17CTNS1	AF112441 Homo sapi	C 784	19	0.3	88459	2	AP001136	AP001136 Homo sapi
C 712	19	0.3	9942	10	AF329470	AF329470 Rattus no	C 785	19	0.3	88712	2	AC097315	AC097315 Mus muscu
C 713	19	0.3	10029	1	AE010387	AE010387 Methanopy	C 786	19	0.3	88823	2	AL1713965	AL1713965 Human DNA
C 714	19	0.3	10029	1	AE014485	AE014485 Rattus no	C 787	19	0.3	90100	2	AC104180	AC104180 Leishmani
C 715	19	0.3	10707	1	AE000099	AE000099 Rhizobium	C 788	19	0.3	92799	2	AC135817	AC135817 Rattus no
C 716	19	0.3	10708	10	RNTRPR2R	XC1677 Rat TRPR2 g	C 789	19	0.3	92956	2	AC138707	AC138707 Homo sapi
C 717	19	0.3	11614	1	AE009582	AE009582 Brucella	C 790	19	0.3	93280	2	AL159174	AL159174 Human DNA
C 718	19	0.3	11688	1	AE009454	AE009454 Brucella	C 791	19	0.3	93280	2	AL159174	AL159174 Human DNA
C 719	19	0.3	12081	10	AL772229	AL772229 Mouse DNA	C 792	19	0.3	93393	4	AC104463	AC104463 Sus scrofa
C 720	19	0.3	12631	1	AE014357	AE014357 Brucella	C 793	19	0.3	95108	2	AC127818	AC127818 Rattus no
C 721	19	0.3	12938	1	AE012470	AE012470 Xanthomon	C 794	19	0.3	95347	9	HS0181122	HS0181122 Homo sapi
C 722	19	0.3	14207	10	AF463762	AF463762 Mus muscu	C 795	19	0.3	95546	9	AC022816	AC022816 Homo sapi

C 796	19	0.3	96009	2	EX545853	869	19	0.3	142114	8	AC073166	AC073166	Oryza sat
C 797	19	0.3	99342	9	HS201D17	870	19	0.3	142560	2	AC021906	AC021906	Homo sapi
C 798	19	0.3	100000	9	AP000215	871	19	0.3	142777	10	AC006404	AC006404	Mus muscu
C 799	19	0.3	104309	9	AL391358	872	19	0.3	142781	2	AC141934	AC141934	Rattus no
C 800	19	0.3	104663	10	AL732560	873	19	0.3	143165	2	AL591114	AL591114	Homo sapi
C 801	19	0.3	105616	10	AL935134	874	19	0.3	143378	5	AC140945	AC140945	Gallus ga
C 802	19	0.3	107226	9	AC010618	875	19	0.3	143805	2	HSB335E1	HSB335E1	Human DNA
C 803	19	0.3	107413	4	AL773560	876	19	0.3	144500	2	AC010231	AC010231	Homo sapi
C 804	19	0.3	107611	8	AC0173560	877	19	0.3	144794	2	AP001182	AP001182	Homo sapi
C 805	19	0.3	108000	9	AP001180	878	19	0.3	145879	10	AL928915	AL928915	Mouse DNA
C 806	19	0.3	109347	9	AL137157	879	19	0.3	145918	9	AC012307	AC012307	Homo sapi
C 807	19	0.3	110000	2	AC009579	880	19	0.3	146153	3	AC007452	AC007452	Drosophi
C 808	19	0.3	110000	2	AC079517	881	19	0.3	147555	2	AC142206	AC142206	Homo sapi
C 809	19	0.3	110000	2	AC096007	882	19	0.3	147727	2	AC011950	AC011950	Homo sapi
C 810	19	0.3	110000	2	AC096315	883	19	0.3	147829	2	AC068806	AC068806	Mus muscu
C 811	19	0.3	110000	2	AC109940	884	19	0.3	149030	9	AC005375	AC005375	Homo sapi
C 812	19	0.3	110000	2	AC110111	885	19	0.3	149087	9	AL157826	AL157826	Homo sapi
C 813	19	0.3	110000	2	AC114711	886	19	0.3	149320	2	OSJN00083	OSJN00083	Oryza sat
C 814	19	0.3	110000	2	AC116152	887	19	0.3	150206	8	HSB305P22	HSB305P22	Human DNA
C 815	19	0.3	110000	2	AC117962	888	19	0.3	151163	9	AC102905	AC102905	Mus muscu
C 816	19	0.3	110000	2	AC118670	889	19	0.3	151839	2	AC024349	AC024349	Homo sapi
C 817	19	0.3	110000	2	AC130075	890	19	0.3	151903	2	AC091756	AC091756	Sus scrof
C 818	19	0.3	110000	2	AC145312	891	19	0.3	152058	4	AC091756	AC091756	Homo sapi
C 819	19	0.3	110000	2	AL354832	892	19	0.3	152209	2	AL591023	AL591023	Homo sapi
C 820	19	0.3	110000	2	AL954350	893	19	0.3	152246	9	AC023473	AC023473	Homo sapi
C 821	19	0.3	110000	2	AL954350	894	19	0.3	152464	2	AC109229	AC109229	Mus muscu
C 822	19	0.3	111106	5	EX663531	895	19	0.3	152827	2	AC143347	AC143347	Homo sapi
C 823	19	0.3	111891	9	AC004797	896	19	0.3	153021	10	AL663042	AL663042	Mouse DNA
C 824	19	0.3	115888	9	AC002468	897	19	0.3	153154	2	AC117770	AC117770	Mus muscu
C 825	19	0.3	116022	9	AL355533	898	19	0.3	153358	2	AC145482	AC145482	Silurana
C 826	19	0.3	116311	9	AL360181	899	19	0.3	153723	9	AC074101	AC074101	Homo sapi
C 827	19	0.3	118495	9	AL590556	900	19	0.3	154262	2	AC127629	AC127629	Rattus no
C 828	19	0.3	118593	9	HS288E11	901	19	0.3	154479	2	BX255900	BX255900	Danio rer
C 829	19	0.3	119066	2	AC026377	902	19	0.3	154983	2	AC018866	AC018866	Homo sapi
C 830	19	0.3	119295	2	AC127427	903	19	0.3	154997	2	AC143903	AC143903	Macaca mu
C 831	19	0.3	120311	10	AC006945	904	19	0.3	155021	9	AC007159	AC007159	Homo sapi
C 832	19	0.3	120613	2	AC111401	905	19	0.3	155278	9	HSB0655C4	HSB0655C4	Human DNA
C 833	19	0.3	123159	9	AC090717	906	19	0.3	155345	9	AC137627	AC137627	Homo sapi
C 834	19	0.3	123022	2	HSAC000384	907	19	0.3	155382	9	AC127520	AC127520	Homo sapi
C 835	19	0.3	123489	9	HSBA27F12	908	19	0.3	155663	3	AC092048	AC092048	Homo sapi
C 836	19	0.3	124289	9	AC010285	909	19	0.3	155670	2	AC012097	AC012097	Drosophi
C 837	19	0.3	124392	9	AL136451	910	19	0.3	155682	2	AL390205	AL390205	Human DNA
C 838	19	0.3	126753	2	AF252826	911	19	0.3	156122	2	AC087296	AC087296	Homo sapi
C 839	19	0.3	126894	9	AL157875	912	19	0.3	156264	9	AC103521	AC103521	Mus muscu
C 840	19	0.3	127431	9	AC027796	913	19	0.3	156339	2	AC132073	AC132073	Bos tauri
C 841	19	0.3	127544	9	CNS01RG4	914	19	0.3	157063	10	AL928950	AL928950	Mouse DNA
C 842	19	0.3	127677	9	AC102945	915	19	0.3	157340	2	AC015652	AC015652	Homo sapi
C 843	19	0.3	127690	2	AC138628	916	19	0.3	157340	2	AC015652	AC015652	Homo sapi
C 844	19	0.3	128090	2	AC108572	917	19	0.3	157519	5	AC146347	AC146347	Pan trogl
C 845	19	0.3	128608	10	AL928638	918	19	0.3	157530	9	AC146347	AC146347	Pan trogl
C 846	19	0.3	129492	9	AC138466	919	19	0.3	157621	2	AC130421	AC130421	Sus scrof
C 847	19	0.3	129506	2	AC136100	920	19	0.3	157654	2	AC091751	AC091751	Rattus no
C 848	19	0.3	130023	2	AC011843	921	19	0.3	157964	2	AC134625	AC134625	Rattus no
C 849	19	0.3	131189	2	AC101756	922	19	0.3	158041	2	AC140688	AC140688	Rattus no
C 850	19	0.3	131353	9	HS608115	923	19	0.3	158156	2	AL357393	AL357393	Homo sapi
C 851	19	0.3	131888	9	HS1057D4	924	19	0.3	158274	2	AC013592	AC013592	Homo sapi
C 852	19	0.3	132150	9	AC005586	925	19	0.3	158307	10	AL592112	AL592112	Mouse DNA
C 853	19	0.3	132336	2	AF215842	926	19	0.3	158338	2	AC024720	AC024720	Homo sapi
C 854	19	0.3	132703	8	CNS608CBT	927	19	0.3	158449	9	AC118689	AC118689	Mus muscu
C 855	19	0.3	133455	2	AC145489	928	19	0.3	158510	9	AC079080	AC079080	Homo sapi
C 856	19	0.3	133698	2	AC143858	929	19	0.3	158920	2	AC135955	AC135955	Canis fam
C 857	19	0.3	133898	2	AC1412917	930	19	0.3	158976	2	AC117771	AC117771	Mus muscu
C 858	19	0.3	135839	9	AC006060	931	19	0.3	159065	3	AC091634	AC091634	Drosophi
C 859	19	0.3	135839	2	AC116945	932	19	0.3	159266	2	AC026097	AC026097	Homo sapi
C 860	19	0.3	136574	2	HS998G20	933	19	0.3	159397	2	AC027060	AC027060	Homo sapi
C 861	19	0.3	136778	2	AC108376	934	19	0.3	159720	2	AC142212	AC142212	Mus muscu
C 862	19	0.3	136956	2	AC137338	935	19	0.3	159720	2	AC068716	AC068716	Homo sapi
C 863	19	0.3	138824	10	AC122519	936	19	0.3	159902	2	AC025571	AC025571	Homo sapi
C 864	19	0.3	139376	9	AC020659	937	19	0.3	159983	2	AC013713	AC013713	Homo sapi
C 865	19	0.3	139376	9	AC020659	938	19	0.3	160074	9	AL105361	AL105361	Homo sapi
C 866	19	0.3	139918	10	AC135564	939	19	0.3	160109	9	AL161727	AL161727	Human DNA
C 867	19	0.3	141017	10	AC126444	940	19	0.3	160396	9	AC107421	AC107421	Homo sapi
C 868	19	0.3	141079	9	HSB218C14	941	19	0.3	160429	2	AC136950	AC136950	Homo sapi

C 942	19	0.3	160452	9	AC110607	Homo sapi
C 943	19	0.3	160677	2	AC125266	Mus muscu
C 944	19	0.3	160785	10	AC009287	Mus muscu
C 945	19	0.3	161025	2	AC101976	Mus muscu
C 946	19	0.3	161181	2	AC116868	Mus muscu
C 947	19	0.3	161218	9	AC011501	Homo sapi
C 948	19	0.3	161223	2	AC124011	Mus muscu
C 949	19	0.3	161251	2	AC147284	Pan trogl
C 950	19	0.3	161309	2	AC100839	Homo sapi
C 951	19	0.3	161506	2	AC092088	Canis fam
C 952	19	0.3	161826	2	AC026157	Homo sapi
C 953	19	0.3	162172	8	AC140026	Medicago
C 954	19	0.3	162213	2	AC061981	Homo sapi
C 955	19	0.3	162377	2	AC142205	Homo sapi
C 956	19	0.3	162471	9	AC067945	Homo sapi
C 957	19	0.3	162573	10	AC125121	Mus muscu
C 958	19	0.3	162877	9	AC018509	Homo sapi
C 959	19	0.3	163073	2	AC118746	Mus muscu
C 960	19	0.3	163108	8	AC137921	Oryza sat
C 961	19	0.3	163283	9	AL355606	Human DNA
C 962	19	0.3	163377	9	AC087714	Homo sapi
C 963	19	0.3	164014	10	AC138116	Mus muscu
C 964	19	0.3	164123	9	AL360078	Human DNA
C 965	19	0.3	164224	2	AC122791	Mus muscu
C 966	19	0.3	164515	2	AC120298	Rattus no
C 967	19	0.3	164692	10	AC127586	Mus muscu
C 968	19	0.3	164831	9	AL590406	Human DNA
C 969	19	0.3	165018	9	AC079173	Homo sapi
C 970	19	0.3	165329	9	AL583850	Human DNA
C 971	19	0.3	165447	9	AC027575	Homo sapi
C 972	19	0.3	166019	2	AC015648	Homo sapi
C 973	19	0.3	166296	2	AC102929	Mus muscu
C 974	19	0.3	166326	2	AC101741	Mus muscu
C 975	19	0.3	166335	9	AC010313	Homo sapi
C 976	19	0.3	166338	9	AC026165	Homo sapi
C 977	19	0.3	166847	9	AC008393	Homo sapi
C 978	19	0.3	166973	9	AL451049	Human DNA
C 979	19	0.3	167205	2	AC023871	Homo sapi
C 980	19	0.3	167273	10	AC083894	Mus muscu
C 981	19	0.3	167794	2	AC146591	Gallitltri
C 982	19	0.3	167852	9	AC072061	Homo sapi
C 983	19	0.3	168030	2	AC091462	Mus muscu
C 984	19	0.3	168133	2	AC140997	Rattus no
C 985	19	0.3	168413	2	AC009877	Homo sapi
C 986	19	0.3	168432	10	AC133574	Mus muscu
C 987	19	0.3	168558	2	BX322549	Danio rer
C 988	19	0.3	168712	10	AC126929	Mus muscu
C 989	19	0.3	168857	2	AC021591	Homo sapi
C 990	19	0.3	169010	10	AL672260	Mouse DNA
C 991	19	0.3	169114	2	AC121128	Mus muscu
C 992	19	0.3	169184	9	AC023398	Homo sapi
C 993	19	0.3	169217	9	AC139149	Homo sapi
C 994	19	0.3	169457	2	AC009216	Drosophila
C 995	19	0.3	170241	2	AC083967	Homo sapi
C 996	19	0.3	170301	2	AC068487	Homo sapi
C 997	19	0.3	170433	2	AC119357	Rattus no
C 998	19	0.3	170758	2	AC004965	Homo sapi
C 999	19	0.3	170772	2	AC016454	Homo sapi
C1000	19	0.3	171042	2	AC120521	Canis fam

ALIGNMENTS

RESULT 1
 AX503782
 LOCUS
 DEFINITION Sequence 12 from Patent WO0226826.
 ACCESSION AX503782
 VERSION AX503782.1 GI:23385966
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

5691 bp DNA
 linear PAT 27-SEP-2002

QY	1	ATGAAAGCATGCGCTGGAAGCTGACCTGCTCTCCCACTCTCATGCTGCGCATG	60
DB	1	ATGAAAGCATGCGCTGGAAGCTGACCTGCTCTCCCACTCTCATGCTGCGCATG	60
QY	61	GGCTGCTCCACCTTGTCTCACTGCGGACGAGCCCGCTGTCAGAGAGAGCGGCTATT	120
DB	61	GGCTGCTCCACCTTGTCTCACTGCGGACGAGCCCGCTGTCAGAGAGAGCGGCTATT	120
QY	121	GTCACATTCGAGAGAGAGCGCGCGAGAGGTTTCAATCACTGCTGCTGATGAGAGACA	180
DB	121	GTCACATTCGAGAGAGAGCGCGCGAGAGGTTTCAATCACTGCTGCTGATGAGAGACA	180
QY	181	GGACACATTTCTGCTGCGGCGCGCTCAATGAGATTACAGAGCTGCTCAGAGAGCTG	240
DB	181	GGACACATTTCTGCTGCGGCGCGCTCAATGAGATTACAGAGCTGCTCAGAGAGCTG	240
QY	241	TTGGTGAAGCATGAGAGAGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	300
DB	241	TTGGTGAAGCATGAGAGAGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	300
QY	301	GTCACAGCTGCTGAG	360
DB	301	GTCACAGCTGCTGAG	360
QY	361	GACTACAG	420
DB	361	GACTACAG	420
QY	421	CTGAGCTGAG	480
DB	421	CTGAGCTGAG	480
QY	481	TCAAGCTGCTGAG	540
DB	481	TCAAGCTGCTGAG	540
QY	541	GACAGCTGCTGAG	600
DB	541	GACAGCTGCTGAG	600
QY	601	AGCCGAACTGAG	660
DB	601	AGCCGAACTGAG	660
QY	661	GAGTGTGCTGAG	720
DB	661	GAGTGTGCTGAG	720
QY	721	GATATCTAT	780
DB	721	GATATCTAT	780

REFERENCE
 AUTHORS
 Gerlach V.L., Macdougall, J.R., Smithson, G., Millet, I., Stone, D.,
 Gunther, E., Ellerman, K., Grosse, W.M., Alsbrook, J.P., Lepley, D.M.,
 Shumke, C.E., Padigaru, M., Kerkula, R., Spyrek, K.A., Leach, M.D. and
 Shimkets, R.A.
 TITLE
 Proteins and nucleic acids encoding same
 JOURNAL
 Patent: WO 0226826-A 12 04-APR-2002;
 Curagen Corporation (US)
 FEATURES
 source
 1. 5691
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 100.0%; Score 5691; DB 6; Length 5691;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 5691; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 781 CCTGAGATGATGTCTCCACGAGCTCCACCAAGAGAGAGATGATACATCCAGCTC 840
Db 781 CCTGAGATGATGTCTCCACGAGCTCCACCAAGAGAGAGATGATACATCCAGCTC 840
QY 841 GTGAGGCTTTGACAG 900
Db 841 GTGAGGCTTTGACAG 900
QY 901 GAGGCGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 960
Db 901 GAGGCGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 960
QY 961 GTGCTTGGCAG 1020
Db 961 GTGCTTGGCAG 1020
QY 1021 AAGGCGCAG 1080
Db 1021 AAGGCGCAG 1080
QY 1081 AAGCAGATTAATGACCGCATTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Db 1081 AAGCAGATTAATGACCGCATTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
QY 1141 CTGAGCCTGGCCTGGCTCAAGGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
Db 1141 CTGAGCCTGGCCTGGCTCAAGGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
QY 1201 GACGATTAATCTTGTGTGGCTTGAATGATGATGATGATGATGATGATGATG 1260
Db 1201 GACGATTAATCTTGTGTGGCTTGAATGATGATGATGATGATGATGATGATG 1260
QY 1261 GAAATTTCCCGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Db 1261 GAAATTTCCCGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
QY 1321 AAGAACCACTCTGTGGCTTTGTGGGACCAAAAGTGGAGAGAGAGAGAGAGAG 1380
Db 1321 AAGAACCACTCTGTGGCTTTGTGGGACCAAAAGTGGAGAGAGAGAGAGAGAG 1380
QY 1381 GATGAG 1440
Db 1381 GATGAG 1440
QY 1441 GTCTCTCGAGATATGGCTTTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
Db 1441 GTCTCTCGAGATATGGCTTTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
QY 1501 CAGCTCAACAG 1560
Db 1501 CAGCTCAACAG 1560
QY 1561 GAGTCAAG 1620
Db 1561 GAGTCAAG 1620
QY 1621 CGGTGTGAGCGGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
Db 1621 CGGTGTGAGCGGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
QY 1681 CTGACGAGTCAATCCCAAGATATCTCGCTCTCAAGTCAAGTCAAGTCAAGTCA 1740
Db 1681 CTGACGAGTCAATCCCAAGATATCTCGCTCTCAAGTCAAGTCAAGTCAAGTCA 1740
QY 1741 ACGTCAATATGTCCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
Db 1741 ACGTCAATATGTCCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
QY 1801 ATGATGAG 1860
Db 1801 ATGATGAG 1860
QY 1861 CCCGAGATCATCAAGAGATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920

Db 1861 CCCGAGATCATCAAGAGATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
QY 1921 GAGACCGAGATGACCTTGGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980
Db 1921 GAGACCGAGATGACCTTGGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980
QY 1981 TGTGTGCTGTCTGT 2040
Db 1981 TGTGTGCTGTCTGT 2040
QY 2041 TGCACCATGAG 2100
Db 2041 TGCACCATGAG 2100
QY 2101 TGT 2160
Db 2101 TGT 2160
QY 2161 AAGCTGAG 2220
Db 2161 AAGCTGAG 2220
QY 2221 CTCAACATTCAG 2280
Db 2221 CTCAACATTCAG 2280
QY 2281 CAGTGCAG 2340
Db 2281 CAGTGCAG 2340
QY 2341 TTGACAGT 2400
Db 2341 TTGACAGT 2400
QY 2401 CTCTCAAGT 2460
Db 2401 CTCTCAAGT 2460
QY 2461 TTGTGATGT 2520
Db 2461 TTGTGATGT 2520
QY 2521 CAG 2580
Db 2521 CAG 2580
QY 2581 ACAGAGATATATCCCGT 2640
Db 2581 ACAGAGATATATCCCGT 2640
QY 2641 GAGAACCTGT 2700
Db 2641 GAGAACCTGT 2700
QY 2701 TGTGAGCCTTTTGT 2760
Db 2701 TGTGAGCCTTTTGT 2760
QY 2761 GCGAAGCCAG 2820
Db 2761 GCGAAGCCAG 2820
QY 2821 TTGATGAG 2880
Db 2821 TTGATGAG 2880
QY 2881 CCCAGCCGAG 2940
Db 2881 CCCAGCCGAG 2940
QY 2941 GCGGAG 3000

Db 2941 GCCGGAAGCAACGTGCTGATGTTGGAAGACGCCCTGCTCTTCCACAGCCGATCT 3000
QY 3001 CCATCTCAATGTTCTGCAACCAACATCCTCAGATGAGTGCTAGAGATGAAGTGTG 3060
Db 3001 CCATCTCAATGTTCTGCAACCAACATCCTCAGATGAGTGCTAGAGATGAAGTGTG 3060
QY 3061 GTGACGTGACAGGAGCCAGATCCAGAGACCTGCTTCACTATGTAATGAAGACCC 3120
Db 3061 GTGACGTGACAGGAGCCAGATCCAGAGACCTGCTTCACTATGTAATGAAGACCC 3120
QY 3121 ACCATCTGCGGATTTGAGCCAGAAATGAGCAATGTCAGTGGAAAACAACCCTGCGGTA 3180
Db 3121 ACCATCTGCGGATTTGAGCCAGAAATGAGCAATGTCAGTGGAAAACAACCCTGCGGTA 3180
QY 3181 TGGGAGACCCCACTGGAACCTCATACAGAACCCCAAGATCCGTCGCAAGATGAGGAG 3240
Db 3181 TGGGAGACCCCACTGGAACCTCATACAGAACCCCAAGATCCGTCGCAAGATGAGGAG 3240
QY 3241 GAGCACATCAATATCTGTAGGTTCTGAACGCTACTAGATGACCTGTCAAGCCCGCC 3300
Db 3241 GAGCACATCAATATCTGTAGGTTCTGAACGCTACTAGATGACCTGTCAAGCCCGCC 3300
QY 3301 CTGCTCTGGGTTCTGACCAACAGTCAAGCTTACCTCAACAGCACTTCACTATCC 3360
Db 3301 CTGCTCTGGGTTCTGACCAACAGTCAAGCTTACCTCAACAGCACTTCACTATCC 3360
QY 3361 ATCTGGAACAAGTCAGTCCCTGCTCATCTCAACAGCACTTCACTATCC 3420
Db 3361 ATCTGGAACAAGTCAGTCCCTGCTCATCTCAACAGCACTTCACTATCC 3420
QY 3421 AACCCGTTGTTGAGGCTTTGCTCTCAAGAAATCTGAGCTCAAGCTTGGACGCC 3480
Db 3421 AACCCGTTGTTGAGGCTTTGCTCTCAAGAAATCTGAGCTCAAGCTTGGACGCC 3480
QY 3481 ATCATCTTAAAGGAGCAAGCACTGATCCGCGCTGAGCTGAGGAGCAAGTGAAGCTGAC 3540
Db 3481 ATCATCTTAAAGGAGCAAGCACTGATCCGCGCTGAGCTGAGGAGCAAGTGAAGCTGAC 3540
QY 3541 TACACTGTGCTGTTGGGAGAAAGCCGTGACCCGTGACCTGTCAAGTGTCCAGCTGTC 3600
Db 3541 TACACTGTGCTGTTGGGAGAAAGCCGTGACCCGTGACCTGTCAAGTGTCCAGCTGTC 3600
QY 3601 TGGGAGTCCCAACCTCATGAGGAGCAAGCAAGTGAAGGACCGGTGAGGATGAG 3660
Db 3601 TGGGAGTCCCAACCTCATGAGGAGCAAGCAAGTGAAGGACCGGTGAGGATGAG 3660
QY 3661 TACTCCCGGGAGTGGTGAATGATGCCCGGACAGCCGCTCAGCTCCGCAATGTC 3720
Db 3661 TACTCCCGGGAGTGGTGAATGATGCCCGGACAGCCGCTCAGCTCCGCAATGTC 3720
QY 3721 AGCATTCGAGTGGCTGGCGGCTCTCATCATTTTCACTGAGCGGTGATGGCCAT 3780
Db 3721 AGCATTCGAGTGGCTGGCGGCTCTCATCATTTTCACTGAGCGGTGATGGCCAT 3780
QY 3781 AAACCAAGTCCCGGAAAGTGAACCTCAAGCTGAAGCGGCTGAGATGCAATGAGCAAC 3840
Db 3781 AAACCAAGTCCCGGAAAGTGAACCTCAAGCTGAAGCGGCTGAGATGCAATGAGCAAC 3840
QY 3841 CTGAGTCCCGTGTGGCTCTGAGTGAAGAAAGCTTTGCGGAGTGAACAGGACATC 3900
Db 3841 CTGAGTCCCGTGTGGCTCTGAGTGAAGAAAGCTTTGCGGAGTGAACAGGACATC 3900
QY 3901 CATGAGTGAACAGTGAACCTGATGAGAGCGGATTCGTTCTGAGCTATAGAACTTAC 3960
Db 3901 CATGAGTGAACAGTGAACCTGATGAGAGCGGATTCGTTCTGAGCTATAGAACTTAC 3960
QY 3961 ACCATGCGGAGTGTCTTCCAGAAATGAAGCAACCTGTCTCCGAGACCTTGAAGTTC 4020
Db 3961 ACCATGCGGAGTGTCTTCCAGAAATGAAGCAACCTGTCTCCGAGACCTTGAAGTTC 4020
QY 4021 CCAGGCTCAAGGAGGCTGTGAGAAAGGCTGAAGCTTTGCGCAAGCTATCAAC 4080
Db 4021 CCAGGCTCAAGGAGGCTGTGAGAAAGGCTGAAGCTTTGCGCAAGCTATCAAC 4080
QY 4081 AACAAAGTGTCTGCTGCTCTTCAATCCGACCGCTTGAAGTCCAGGCTGCTCCATG 4140
Db 4081 AACAAAGTGTCTGCTGCTCTTCAATCCGACCGCTTGAAGTCCAGGCTGCTCCATG 4140
QY 4141 CCGGACCGTGAAGTGGCTCACTCATCATGACCGTGTGCAAGAGCAAGCTGAGTAC 4200
Db 4141 CCGGACCGTGAAGTGGCTCACTCATCATGACCGTGTGCAAGAGCAAGCTGAGTAC 4200
QY 4201 GCACTGATGTGCTGAGAGAGTGTGTCGACCACTCATTTGACAAAGAACTGAGAGCAG 4260
Db 4201 GCACTGATGTGCTGAGAGAGTGTGTCGACCACTCATTTGACAAAGAACTGAGAGCAG 4260
QY 4261 AACCAACCTTAACTGCTGCTGAGAGAGCTGAGTGAAGTGAAGATGCTGCAAT 4320
Db 4261 AACCAACCTTAACTGCTGCTGAGAGAGCTGAGTGAAGTGAAGATGCTGCAAT 4320
QY 4321 TGGTTTACTTTCTCTCTTCAAGTTCTTCAAGAGTGTGCTGGGAGCCCTCTTCTCC 4380
Db 4321 TGGTTTACTTTCTCTCTTCAAGTTCTTCAAGAGTGTGCTGGGAGCCCTCTTCTCC 4380
QY 4381 CTGTTCTGTGCAATCAAGAGAGATGAGAGAGTGTGCTGAGGAGCTGAGTGAAGATGCTGCAAT 4440
Db 4381 CTGTTCTGTGCAATCAAGAGAGATGAGAGAGTGTGCTGAGGAGCTGAGTGAAGATGCTGCAAT 4440
QY 4441 GCCCGTACTCTTGAAGGAGCAAGCTCATCCGTCAGAGATTGACTACAAACCTTG 4500
Db 4441 GCCCGTACTCTTGAAGGAGCAAGCTCATCCGTCAGAGATTGACTACAAACCTTG 4500
QY 4501 GTCCTGAGTGTGAGAGGAGCAAGTCAAGTCCAGAGGAGTCCAGTAAAGTCTC 4560
Db 4501 GTCCTGAGTGTGAGAGGAGCAAGTCAAGTCCAGAGGAGTCCAGTAAAGTCTC 4560
QY 4561 AACTGTGACACATCACTGAGTCAAGAGAAAGATTGTGATGCTTCAAGATGTG 4620
Db 4561 AACTGTGACACATCACTGAGTCAAGAGAAAGATTGTGATGCTTCAAGATGTG 4620
QY 4621 CCTGTCCACCGGCCAAAGCTGAGATGAGTCTGAGTGGCGCAAGAAAGTGG 4680
Db 4621 CCTGTCCACCGGCCAAAGCTGAGATGAGTCTGAGTGGCGCAAGAAAGTGG 4680
QY 4681 GCAAGGATGATCTTGAAGATGAAGACATCACCAACAAATTGAGATGTGAAGCGA 4740
Db 4681 GCAAGGATGATCTTGAAGATGAAGACATCACCAACAAATTGAGATGTGAAGCGA 4740
QY 4741 CTGAAACACTGGCCCACTACAGGTGCCAATGTTCCGTGGGCAATTAATGCCAAG 4800
Db 4741 CTGAAACACTGGCCCACTACAGGTGCCAATGTTCCGTGGGCAATTAATGCCAAG 4800
QY 4801 CAGGTGACAGCTATPACAGTGAACAACCTCAAGCTTCCAGACCTCAAGCAAGTAA 4860
Db 4801 CAGGTGACAGCTATPACAGTGAACAACCTCAAGCTTCCAGACCTCAAGCAAGTAA 4860
QY 4861 TATGAAAATGATCCGTGACACCGGAGCCCGGACAGCTCCGCTACAGGCACTTATG 4920
Db 4861 TATGAAAATGATCCGTGACACCGGAGCCCGGACAGCTCCGCTACAGGCACTTATG 4920
QY 4921 ATCACTCTGACCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4980
Db 4921 ATCACTCTGACCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4980
QY 4981 GAGACCAAGAGAGGAGGAGCCGGGAGCAAGATGTTCTGAAATCTACCTGACCGA 5040
Db 4981 GAGACCAAGAGAGGAGGAGCCGGGAGCAAGATGTTCTGAAATCTACCTGACCGA 5040
QY 5041 CTCTGACCACTPAAAGGACACCTGAGAGATTGAGATGACTCTTGAAGCAATCTTC 5100
Db 5041 CTCTGACCACTPAAAGGACACCTGAGAGATTGAGATGACTCTTGAAGCAATCTTC 5100
QY 5101 AGCAGGCAACCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5160
Db 5101 AGCAGGCAACCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5160

QY 5161 GATGAGCAGGCTGATTAACATGGCATTGATGACCCCGACGCTCCGCATACCTTGGAAGAC 5220
DB 5161 GATGAGCAGGCTGATTAACATGGCATTGATGACCCCGACGCTCCGCATACCTTGGAAGAC 5220
QY 5221 AATTGCTGCCCCGAGCTTTGGGTCACATGATGACAGAACCGGAGTTGTGTTGAC 5280
DB 5221 AATTGCTGCCCCGAGCTTTGGGTCACATGATGACAGAACCGGAGTTGTGTTGAC 5280
QY 5281 ATCCATTAAGAACACATCAACAGACGCGCTGCTCTGTGTGTGCTGACCTTCATGAC 5340
DB 5281 ATCCATTAAGAACACATCAACAGACGCGCTGCTCTGTGTGTGCTGACCTTCATGAC 5340
QY 5341 TCTTGTCTCAGTGTAGAGACACCGGCTGGGCAAGACCTCCGCTCCAAACAGCTGTGTAT 5400
DB 5341 TCTTGTCTCAGTGTAGAGACACCGGCTGGGCAAGACCTCCGCTCCAAACAGCTGTGTAT 5400
QY 5401 GCCAAGGACATCCCAAGCTACCAAGATTGGGTGGAGAGTATTACTGACATAGGGAAG 5460
DB 5401 GCCAAGGACATCCCAAGCTACCAAGATTGGGTGGAGAGTATTACTGACATAGGGAAG 5460
QY 5461 ATGCCAGCCTACGACGACCAAGACATGACGATACCTGTGCTGACGATCCCGGATGAC 5520
DB 5461 ATGCCAGCCTACGACGACCAAGACATGACGATACCTGTGCTGACGATCCCGGATGAC 5520
QY 5521 ATGAATGAGTCAACACCATGATGACACTCTCAGAGATCTTCTCTATGTGGCCAAATAC 5580
DB 5521 ATGAATGAGTCAACACCATGATGACACTCTCAGAGATCTTCTCTATGTGGCCAAATAC 5580
QY 5581 AGCAGAGAGATCTTGGACCTCTGACACGATGACCAAGTGTGGGAAGCAAACTGGCC 5640
DB 5581 AGCAGAGAGATCTTGGACCTCTGACACGATGACCAAGTGTGGGAAGCAAACTGGCC 5640
QY 5641 TACAACTAGAACAAAGTCATTAACCTCATGAGCTTAGACAGCTGAATTA 5691
DB 5641 TACAACTAGAACAAAGTCATTAACCTCATGAGCTTAGACAGCTGAATTA 5691

RESULT 2
AX492982 6367 bp DNA linear PAT 26-SEP-2002
LOCUS Sequence 19 from Patent WO02059312.
DEFINITION AX492982
ACCESSION AX492982.1 GI:23338637
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
1 Kallik, D.A., Lee, S., Xu, Y., Yao, M.G., Yue, H., Bandman, O.B.,
Buford, N., Gandhi, A.R., Graul, R.C., Lai, P.G., Lu, D.A., Lu, Y.,
Tang, T.Y., Duggan, B.M., Gietzen, R.J., Hillman, J.L., Honchell, C.D.,
Ramkumar, J., Walla, N.K. and Warren, B.A.
Cell adhesion proteins
Patent: WO 02059312-A 19 01-AUG-2002;
JOURNAL INCTE GENOMICS INC (US)
FEATURES
source
1. 6367
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyle ID No: 7156379CB."

ORIGIN
Query Match 97.2%; Score 5532; DB 6; Length 6367;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 5682; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGAAGCAGTCCCTGGAATGACCTGCTCTCCCACTCTCATGTTGGGATG 60
DB 566 ATGAAGCAGTCCCTGGAATGACCTGCTCTCCCACTCTCATGTTGGGATG 625
QY 61 GGCCTCTCCTTGTGTCAACCGGACGACGCGCTGTCCAGAGACGCGTCA 120

DB 626 GGCCTCTCCTTGTGTCAACCGGACGACGCGCTGTCCAGAGACGCGTCA 685
QY 121 GTCACTTCCAGAGAGACCCCGAGGGTTTCAATCACTGTGTGTGATGAGAGACA 180
DB 686 GTCACTTCCAGAGAGACCCCGAGGGTTTCAATCACTGTGTGTGATGAGAGACA 745
QY 181 GGAACATTTACTTGGGGGCGCTCATGTGATTTTCAAGCTCTCAGAGACCTGAAGT 240
DB 746 GGAACATTTACTTGGGGGCGCTCATGTGATTTTCAAGCTCTCAGAGACCTGAAGT 805
QY 241 TTGTGAGCATGAGACAGGCGCGACAGAGACAACCCCAAGTGTACCCAGCCGATC 300
DB 806 TTGTGAGCATGAGACAGGCGCGACAGAGACAACCCCAAGTGTACCCAGCCGATC 865
QY 301 GTCCAGACCTGCAATGACCCCTTGAACACCAACCAATGTCAACAAATGCTCTCAT 360
DB 866 GTCCAGACCTGCAATGACCCCTTGAACACCAACCAATGTCAACAAATGCTCTCAT 925
QY 361 GACTACAGAGAGAACAGGCTGATGCTGTGGAGGCTGTACCAAGCATCTGAAGCTG 420
DB 926 GACTACAGAGAGAACAGGCTGATGCTGTGGAGGCTGTACCAAGCATCTGAAGCTG 985
QY 421 CTGAGGCTGAGAGACCTTTTCAAGCTGGGGAGCCTTATCATTAAGAGAGACATAT 480
DB 986 CTGAGGCTGAGAGACCTTTTCAAGCTGGGGAGCCTTATCATTAAGAGAGACATAT 1045
QY 481 TCAGGTCACAGAGAGGCGCTCAGTCTTGTGAGTGTCTCTCAAGACCTGAT 540
DB 1046 TCAGGTCACAGAGAGGCGCTCAGTCTTGTGAGTGTCTCTCAAGACCTGAT 1105
QY 541 GACAAAGCTGTTCAATGCCAGCGCACTGAGTGGAAAGCCGAGATTTTCCACATCT 600
DB 1106 GACAAAGCTGTTCAATGCCAGCGCACTGAGTGGAAAGCCGAGATTTTCCACATCT 1165
QY 601 AGCCGGAACCTGACCAAGAACCTGAGGGGAGTGGCATGTCGAGTACGTCCTCAT 660
DB 1166 AGCCGGAACCTGACCAAGAACCTGAGGGGAGTGGCATGTCGAGTACGTCCTCAT 1225
QY 661 GAGTTCGTGGCTCGATGATTAAGATCCCTTGGACACCTTACCATCATCTGACT 720
DB 1226 GAGTTCGTGGCTCGATGATTAAGATCCCTTGGACACCTTACCATCATCTGACT 1285
QY 721 GATATCTACTATGTCTATGCTTTTGTAGAGTGGCACTTGTCTATTTTGGCCCTCA 780
DB 1286 GATATCTACTATGTCTATGCTTTTGTAGAGTGGCACTTGTCTATTTTGGCCCTCA 1345
QY 781 CCTGAGATGCTGTCTCAACAGGCTCAACCAACCAAGAGCAGGTATATATCAAGCTC 840
DB 1346 CCTGAGATGCTGTCTCAACAGGCTCAACCAACCAAGAGCAGGTATATATCAAGCTC 1405
QY 841 GTGAGGCTTTGCAAGAGAGACACGCTTCAACTCTATATATAGAGTGGCTGCT 900
DB 1406 GTGAGGCTTTGCAAGAGAGACACGCTTCAACTCTATATATAGAGTGGCTGCT 1465
QY 901 GAGCGAGTGGGTGAGTACCGCTGTCAGGCTGCTCACTCTCAAAAGCGGGGCGC 960
DB 1466 GAGCGAGTGGGTGAGTACCGCTGTCAGGCTGCTCACTCTCAAAAGCGGGGCGC 1525
QY 961 GTGCTTGGCAGAGACCTTGGATGATCAATGATGATGATGATGATGATGATGAT 1020
DB 1526 GTGCTTGGCAGAGACCTTGGATGATCAATGATGATGATGATGATGATGATGAT 1585
QY 1021 AAGGGCCAGAGGGAATGAATCCCTGATGATGATGATGATGATGATGATGATGAT 1080
DB 1586 AAGGGCCAGAGGGAATGAATCCCTGATGATGATGATGATGATGATGATGATGAT 1645
QY 1081 AAGCAGATTAATGACCGCATTAAGAGCGGCTGCACTGTTTACCGGGGCAAGGCA 1140
DB 1646 AAGCAGATTAATGACCGCATTAAGAGCGGCTGCACTGTTTACCGGGGCAAGGCA 1705
QY 1141 CTGAGCTGCGCTGCTCAAGGTGAAGGACATCCCTGACAGGAGTGCCTTAAACAT 1200

Db 1706 CTGGACCTGGCCCTGGCTCAAGGTAAGGACATCCCTCTGACGAGTGCCTTTAAACATT 1765
QY 1201 GACGATAACTTTGTGTGGCTTGGACATGAATGTCTCCCTGGAGTGTCCGACATGTGTGGT 1260
Db 1766 GACGATAACTTTGTGTGGCTTGGACATGAATGTCTCCCTGGAGTGTCCGACATGTGTGGT 1825
QY 1261 GGAATTCCTGTCTTCAAGAGGACAGGAGCCGATAGAGTGTGTATGTATGTATGTATG 1320
Db 1826 GGAATTCCTGTCTTCAAGAGGACAGGAGCCGATAGAGTGTGTATGTATGTATGTATG 1885
QY 1321 AAGAACCACTCTGTGGCTTTGTGGACACAAAAGTGGCAAGGTGAAGAATCCGGGTG 1380
Db 1886 AAGAACCACTCTGTGGCTTTGTGGACACAAAAGTGGCAAGGTGAAGAATCCGGGTG 1945
QY 1381 GATGAGCCCAAGGGGCAACGCCCTCCAGTATGAGACGTTGAGTGTGGACCCCGGCCA 1440
Db 1946 GATGAGCCCAAGGGGCAACGCCCTCCAGTATGAGACGTTGAGTGTGGACCCCGGCCA 2005
QY 1441 GTCTCCGGGATATGGCTTTCTCCAGAGACCAAGCAACTTACATCATGTCAAGAGG 1500
Db 2006 GTCTCCGGGATATGGCTTTCTCCAGAGACCAAGCAACTTACATCATGTCAAGAGG 2065
QY 1501 CAGCTCAACAGAGTCCCTGTGTGAGTCTGTGTCTAGTATCAAGCTGCGCGAGTGCCTT 1560
Db 2066 CAGCTCAACAGAGTCCCTGTGTGAGTCTGTGTCTAGTATCAAGCTGCGCGAGTGCCTT 2125
QY 1561 GGCTCAGAGCCACCCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1620
Db 2126 GGCTCAGAGCCACCCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2185
QY 1621 CGGTGTAGCGGT 1680
Db 2186 CGGTGTAGCGGT 2245
QY 1681 CTGACGGTCACTCCCAACATATCTCGTCTCAGTACAACTGTGTGTGTGTGTGTGTGTGT 1740
Db 2246 CTGACGGTCACTCCCAACATATCTCGTCTCAGTACAACTGTGTGTGTGTGTGTGTGTGT 2305
QY 1741 ACGTACAATGTCCCGAGGT 1800
Db 2306 ACGTACAATGTCCCGAGGT 2365
QY 1801 ATGTATGGGT 1860
Db 2366 ATGTATGGGT 2425
QY 1861 CCCCCGATCACTCAAGAAATGGGGACCAACATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1920
Db 2426 CCCCCGATCACTCAAGAAATGGGGACCAACATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2485
QY 1921 GAGACCCGCGATGACCTTGT 1980
Db 2486 GAGACCCGCGATGACCTTGT 2545
QY 1981 TCGTGCTGT 2040
Db 2546 TCGTGCTGT 2605
QY 2041 TGCACCCATGACCCCAAGACTGT 2100
Db 2606 TGCACCCATGACCCCAAGACTGT 2665
QY 2101 TGCCCCAGCTGT 2160
Db 2666 TGCCCCAGCTGT 2725
QY 2161 AGCTGAAAGCCAAAGAACTTCCCAAGCCCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2220
Db 2726 AGCTGAAAGCCAAAGAACTTCCCAAGCCCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2785
QY 2221 CTCAACATTCAGGGCAGGACAGACGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2280
Db 2786 CTCAACATTCAGGGCAGGACAGACGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2845

QY 2281 CAGTGCAGAAACACCTCTTATTCCTATGAAAGGATGGAGATCAACAACCTGCCCCGTGGAG 2340
Db 2846 CAGTGCAGAAACACCTCTTATTCCTATGAAAGGATGGAGATCAACAACCTGCCCCGTGGAG 2905
QY 2341 TTGACAGT 2400
Db 2906 TTGACAGT 2965
QY 2401 CTCTACAAGT 2460
Db 2966 CTCTACAAGT 3025
QY 2461 TTGCGATGT 2520
Db 3026 TTGCGATGT 3085
QY 2521 CAGGAGAGCCCAAGT 2580
Db 3086 CAGGAGAGCCCAAGT 3145
QY 2581 ACAGAGATTAATCCCGGT 2640
Db 3146 ACAGAGATTAATCCCGGT 3205
QY 2641 GAGAACTGT 2700
Db 3206 GAGAACTGT 3265
QY 2701 TGACGCTTTAGT 2760
Db 3266 TGACGCTTTAGT 3325
QY 2761 GCCAAGCCCAAGCAGATGT 2820
Db 3326 GCCAAGCCCAAGCAGATGT 3385
QY 2821 TTCAATGGCGGT 2880
Db 3386 TTCAATGGCGGT 3445
QY 2881 CCGACGCGGGGCCCATGT 2940
Db 3446 CCGACGCGGGGCCCATGT 3505
QY 2941 GCCGGAGACCAAGT 3000
Db 3506 GCCGGAGACCAAGT 3565
QY 3001 CCATCTCACTATGT 3060
Db 3566 CCATCTCACTATGT 3625
QY 3061 GTGCAAGT 3120
Db 3626 GTGCAAGT 3685
QY 3121 ACCATGT 3180
Db 3686 ACCATGT 3745
QY 3181 TGGGGGACCCCACTGT 3240
Db 3746 TGGGGGACCCCACTGT 3805
QY 3241 GAGCACTCAATATGT 3300
Db 3806 GAGCACTCAATATGT 3865
QY 3301 CTGCGCTGT 3360
Db 3866 CTGCGCTGT 3925

QY 3361 ATCTGACACAGTCGAGTCCCTGCTCATCTCAACAAACCAACTTCACTTATCTATCCC 3420
Db 3326 ATCTGACACAGTCGAGTCCCTGCTCATCTCAACAAACCAACTTCACTTATCTATCCC 3985
QY 3421 AACCCGGTGTGAGGCTTTGCTGCTCAGGAATCCTGAGGCTCAAGCTGGACAGCCC 3480
Db 3986 AACCCGGTGTGAGGCTTTGCTGCTCAGGAATCCTGAGGCTCAAGCTGGACAGCCC 4045
QY 3481 ATCATCTTAAAGGCGCAAGAACTGTATCCGCTGTGCTGAGGCGCAACGTGAAGCTGAAC 3540
Db 4046 ATCATCTTAAAGGCGCAAGAACTGTATCCGCTGTGCTGAGGCGCAACGTGAAGCTGAAC 4105
QY 3541 TACACTGTGCTGTTGGGAGAGAACCGTGCACCGTGCACCGTGCATGATGTCCAGCTGCTC 3600
Db 4106 TACACTGTGCTGTTGGGAGAGAACCGTGCACCGTGCACCGTGCATGATGTCCAGCTGCTC 4165
QY 3601 TGGCAGTCCCGCAACTCATCGCAGGCAAGAGTGAATGGCCCGGTGTGAGTGCATGAG 3660
Db 4166 TGGCAGTCCCGCAACTCATCGCAGGCAAGAGTGAATGGCCCGGTGTGAGTGCATGAG 4225
QY 3661 TACTTCCCGGGAGTGTGTATCATTTGCCCGGACAGCCCGCTCAGCTGCCCCGCTATGTC 3720
Db 4226 TACTTCCCGGGAGTGTGTATCATTTGCCCGGACAGCCCGCTCAGCTGCCCCGCTATGTC 4285
QY 3721 AGCATGCGAGTGGCTGGGCGGCTCTCATCATTTTCATGCGCGCTCATTTGCTAT 3780
Db 4286 AGCATGCGAGTGGCTGGGCGGCTCTCATCATTTTCATGCGCGCTCATTTGCTAT 4345
QY 3781 AAACGCAAGTCCGCGAAAGTGAAGCTCAAGCTGAAGCGCTGACAGATGACATGACAC 3840
Db 4346 AAACGCAAGTCCGCGAAAGTGAAGCTCAAGCTGAAGCGCTGACAGATGACATGACAC 4405
QY 3841 CTGAGTGTCCGTTGGGCGCTGAGTGAAGGAGCTTTGCGAGCTGACAGACGACATC 3900
Db 4406 CTGAGTGTCCGTTGGGCGCTGAGTGAAGGAGCTTTGCGAGCTGACAGACGACATC 4465
QY 3901 CATGAGCTGACCAAGTGAAGCTGATGAGCGGAGATTCGTTCTGACATTAAGACTTAC 3960
Db 4466 CATGAGCTGACCAAGTGAAGCTGATGAGCGGAGATTCGTTCTGACATTAAGACTTAC 4525
QY 3961 ACCATGCGGAGTGTGCTTCCAGAAATGGAACCAACCTGTCGCGAGACCTTGAAGTTC 4020
Db 4526 ACCATGCGGAGTGTGCTTCCAGAAATGGAACCAACCTGTCGCGAGACCTTGAAGTTC 4585
QY 4021 CCGGCTACCGGAGAGAGCTGTGAGAAAGGCTTGAAGCTTTTCCGCCACTCATCAAC 4080
Db 4586 CCGGCTACCGGAGAGAGCTGTGAGAAAGGCTTGAAGCTTTTCCGCCACTCATCAAC 4645
QY 4081 AACAAAGTGTGCTGCTCTTCAATCCGACGCTTGAAGTCCAGCGTATTCATCATG 4140
Db 4646 AACAAAGTGTGCTGCTCTTCAATCCGACGCTTGAAGTCCAGCGTATTCATCATG 4705
QY 4141 CGCGACCGTGAAGAGTGGCTCACTCATGACCGTGCAGAGACCAAGCTGAGTAC 4200
Db 4706 CGCGACCGTGAAGAGTGGCTCACTCATGACCGTGCAGAGACCAAGCTGAGTAC 4765
QY 4201 GCGACTGATGTGCTGAAGCAAGCTGTGGCGCACTCATTTGAAGAACTTGAAGACAG 4260
Db 4766 GCGACTGATGTGCTGAAGCAAGCTGTGGCGCACTCATTTGAAGAACTTGAAGACAG 4825
QY 4261 AACCAACCTAAGCTGTGCTGCTGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGT 4320
Db 4826 AACCAACCTAAGCTGTGCTGCTGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGT 4885
QY 4321 TGGTTTACTTCTCTCTCAAGTTCCTCAAGAGTGTGCTGGAGAGCTCTTCTTCC 4380
Db 4886 TGGTTTACTTCTCTCTCAAGTTCCTCAAGAGTGTGCTGGAGAGCTCTTCTTCC 4945
QY 4381 CTGTTCTGTCGATGAAGAGAGATGAGAAAGGCGCCATTGAGCGCATCAAGGAGAG 4440
Db 4946 CTGTTCTGTCGATGAAGAGAGATGAGAAAGGCGCCATTGAGCGCATCAAGGAGAG 5005
QY 4441 GCCCGCTACTCTTGAAGAGAGCAAGCTCATCCGCGACAGATTGACTACAAACCTTG 4500

Db 5006 GCCCGCTACTCTTGAAGAGAGCAAGCTCATCCGCGACAGATTGACTACAAACCTTG 5065
QY 4501 GTCTGAGCTGTGTAGGCCCAAGCAATGCAACAGCCCGAGTCCAGTAAAGTCTTC 4560
Db 5066 GTCTGAGCTGTGTAGGCCCAAGCAATGCAACAGCCCGAGTCCAGTAAAGTCTTC 5125
QY 4561 AACTGTGACCATATCACTCAAGTGAAGAGAAATTCGTGATGCCATCTTAAGATGTG 4620
Db 5126 AACTGTGACCATATCACTCAAGTGAAGAGAAATTCGTGATGCCATCTTAAGATGTG 5185
QY 4621 CTTGCTCCACCGGCGCAAGCTGCAGATATGATCTGAGTGGCGCAAGAAAGTGGG 4680
Db 5186 CTTGCTCCACCGGCGCAAGCTGCAGATATGATCTGAGTGGCGCAAGAAAGTGGG 5245
QY 4681 GCAAGATGATTTGACAGATGAAGACATCAACCAAGATTGAGAAATGATGAGAGGA 4740
Db 5246 GCAAGATGATTTGACAGATGAAGACATCAACCAAGATTGAGAAATGATGAGAGGA 5305
QY 4741 CTGAACACACTGGCCCACTACCAAGTGCAGATGGTCCGTGTGGCATTTAGTGTCAAG 4800
Db 5306 CTGAACACACTGGCCCACTACCAAGTGCAGATGGTCCGTGTGGCATTTAGTGTCAAG 5365
QY 4801 CAGGTGACAGCTTATACGACGTGAACCACTCCACCGTCTCAGGACCTTAA 4860
Db 5366 CAGGTGACAGCTTATACGACGTGAACCACTCCACCGTCTCAGGACCTTAA 5425
QY 4861 TATGAAACATGATCCGATGAACGGGCGAGCCCGGACAGCTCCGCTCAGGACCTTATG 4920
Db 5426 TATGAAACATGATCCGATGAACGGGCGAGCCCGGACAGCTCCGCTCAGGACCTTATG 5485
QY 4921 ATCACTCTGACCTTGAAGATGAGTCAAGATGTGGCACTTATGTAAGAACAGAC 4980
Db 5486 ATCACTCTGACCTTGAAGATGAGTCAAGATGTGGCACTTATGTAAGAACAGAC 5545
QY 4981 GGAAGACAGAAAGAGAGGAGCCGGGGAGCAAGATGTGTGAAATCACTGACCGCA 5040
Db 5546 GGAAGACAGAAAGAGAGGAGCCGGGGAGCAAGATGTGTGAAATCACTGACCGCA 5605
QY 5041 CTCCTGGCCACTTAAGGGCACTGACAGAAATTTGATGATGACTTCTTGAAGCAATCTTC 5100
Db 5606 CTCCTGGCCACTTAAGGGCACTGACAGAAATTTGATGATGACTTCTTGAAGCAATCTTC 5665
QY 5101 AGCAGGCAACCGTGTGCTTGGCCCTGCGCCCTGACATCAAGTATGATGATCTCTG 5160
Db 5666 AGCAGGCAACCGTGTGCTTGGCCCTGCGCCCTGACATCAAGTATGATGATCTCTG 5725
QY 5161 GATGAGCAGCTGATTAACATGATGATGATGATGATGATGATGATGATGATGATGATG 5220
Db 5726 GATGAGCAGCTGATTAACATGATGATGATGATGATGATGATGATGATGATGATGATG 5785
QY 5221 AATTGCTGCGCTTGAAGTGTGCTGATCAATGATCAAGAACCCGCAAGTTGTGTTGAC 5280
Db 5786 AATTGCTGCGCTTGAAGTGTGCTGATCAATGATCAAGAACCCGCAAGTTGTGTTGAC 5845
QY 5281 ATTCATTAAGACAGATCAAGACGCTGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5340
Db 5846 ATTCATTAAGACAGATCAAGACGCTGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5905
QY 5341 TCTTGTCTCAAGTCAAGAGACCGGCTGGGCAAGACTGCGCTTCAAGAAAGTGTGTAT 5400
Db 5906 TCTTGTCTCAAGTCAAGAGACCGGCTGGGCAAGACTGCGCTTCAAGAAAGTGTGTAT 5965
QY 5401 GCCAAGACATCCCGGCTTACAGAAATTTGGAGAGGATTTACTCAGACATTAAGGAG 5460
Db 5966 GCCAAGACATCCCGGCTTACAGAAATTTGGAGAGGATTTACTCAGACATTAAGGAG 6025
QY 5461 ATGCGACCATGACGACCAAGACATGAAGCGATACCTGTGAGAGATCCCGGATGAC 5520
Db 6026 ATGCGACCATGACGACCAAGACATGAAGCGATACCTGTGAGAGATCCCGGATGAC 6085
QY 5521 ATGAATGATTAACACATGAGTGAATCTCAGAGATCTTCTGATGTGGGCAATATC 5580

Db 6086 ATGATGAGTTCAACACCATGAGTGCACCTCCAGAGATCTTCTCCTGATGTGGGCAAAATAC 6145

Qy 5581 AGCAGGAGAGATCCTTGGACCTCCTGGAGCCAGATGATACCGATGTGGGAGACAAACTGGCC 5640

Db 6146 AGCAGGAGAGATCCTTGGACCTCCTGGAGCCAGACGATGTGGGAGACAAAACTGGCC 6205

Qy 5641 TACCAACTGACCAAGATATAAACCCTCATAGCTTAAACAGCTGA 5685

Db 6206 TACCAACTGACCAAGATATAAACCCTCATAGCTTAAACAGCTGA 6250

RESULT 3	5886 bp	linear	PRI 10-MAY-2002
AB046770			
LOCUS	AB046770		
DEFINITION	Homo sapiens mRNA for KIAA1550 protein, partial cds.		
ACCESSION	AB046770		
VERSION	AB046770.2	GI:20521937	
KEYWORDS			
SOURCE	Homo sapiens (human)		

REFERENCE
AUTHORS
TITLE
JOURNAL
DNA Res. 7 (4), 273-281 (2000)

PIUMED 10997877 2 (bases 1 to 5986)
REFERENCE Ohara,O., Nagase,T. and Kikuno,R.
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (03-AUG-2000) Osamu Ohara, Kazusa DNA Research Institute,
Department of Human Gene Research, 1532-3, Yama, Kisarazu, Chiba
293-0842, Japan (E-mail:cdnaifco@kazusa.or.jp,
URL:htp://www.kazusa.or.jp/hnqe, Tel:81-438-55-3913,
Fax:81-438-55-3914)
COMMENT On May 9, 2002 this sequence version replaced gi:10047164.
FEATURES
source Location/Qualifiers
1..5986

gene
CDS

[illegible]

ORIGIN
LECKEGTEMPHAGGHVVCVRVICVCMHICVCCVCEFIYKQAGMAVGSAGSMRCVCLC
EVCVHVCCVCSVCIVSYTSKQADQ"

Query Match		39.3%	Score 2237	DB 9	Length 5986	
Best Local Similarity		100.0%	Ped. No. 0			
Matches 2237		Conservative	0	Mismatches	0	Indels
					0	Gaps
					0	
QY	1638	GGAGCCCCCGCAGGTTTGCCTTCGAGAGTGAAGCACTGTGTCCGGCTGACGGTCCATCCCA	1637			
DB	1	GGAGCCCCCGCAGGTTTGCCTTCGAGAGTGAAGCACTGTGTCCGGCTGACGGTCCATCCCA	60			
QY	1698	CAATATCTCCGTCCTCAGTACAAACGGCTGCTGAGTCTCTGAGACGTACATGTCCCGA	1757			
DB	61	CAATATCTCCGTCCTCAGTACAAACGGCTGCTGAGTCTCTGAGACGTACATGTCCCGA	120			
QY	1758	GCTGTCACTGGGTCACTGCACTTTGAGAACCTGTCAAGATGGAATGGGCTGTCTGT	1817			
DB	121	GCTGTCACTGGGTCACTGCACTTTGAGAACCTGTCAAGATGGAATGGGCTGTCTGT	180			
QY	1818	GGGCAATCAGATCCAGAGCTACTCCCTGAGCCAAAGAGTGGCCCGGATCATCAGA	1877			
DB	181	GGGCAATCAGATCCAGAGCTACTCCCTGAGCCAAAGAGTGGCCCGGATCATCAGA	240			
QY	1878	GAATGGGAGACCACTATGTCTGTAAGCTTCAAGTCAATCAAGAGAACCCGATGACTT	1937			
DB	241	GAATGGGAGACCACTATGTCTGTAAGCTTCAAGTCAATCAAGAGAACCCGATGACTT	300			
QY	1938	CGCCAGACACAGCTTGTGCTTCTACATGTAGAGGTCCAAATTCGAGCTGTCTCGT	1937			
DB	301	CGCCAGACACAGCTTGTGCTTCTACATGTAGAGGTCCAAATTCGAGCTGTCTCGT	360			
QY	1998	GGAGAGTCCATACCGCTGCCACTGGTGTAAATACCGGATGTCTGCACCCATGCCCA	2057			
DB	361	GGAGAGTCCATACCGCTGCCACTGGTGTAAATACCGGATGTCTGCACCCATGCCCA	420			
QY	2058	GACCTGCTCCTTCAGAGAAAGGCGAGAGTCCCGAGGACTGCCCGCCAGCTGTGCG	2117			
DB	421	GACCTGCTCCTTCAGAGAAAGGCGAGAGTCCCGAGGACTGCCCGCCAGCTGTGCG	480			
QY	2118	AGTGAACAAGATCCTGTGTCCCGTGAGGTGATCAAGCTTATCAAGCTGAAAGCCAAAGA	2177			
DB	481	AGTGAACAAGATCCTGTGTCCCGTGAGGTGATCAAGCTTATCAAGCTGAAAGCCAAAGA	540			
QY	2178	CCTCCCCAGGCCCAAGTGTGGGACGGGTGTAAGATGATCTCTCAATTCAAGGGCAG	2237			
DB	541	CCTCCCCAGGCCCAAGTGTGGGACGGGTGTAAGATGATCTCTCAATTCAAGGGCAG	600			
QY	2238	CGAGCAGGAGTGCCTCGCCTTGCGCTTCAACAGCTCAAGCTGATCAGTGCAGAAACCTC	2237			
DB	601	CGAGCAGGAGTGCCTCGCCTTGCGCTTCAACAGCTCAGCTGATCAGTGCAGAAACCTC	660			
QY	2298	TTATTTCCATAAGAGGATGAGATCAACAACCTCCCGTGAAGTGAACAGTGTGGAA	2357			
DB	661	TTATTTCCATAAGAGGATGAGATCAACAACCTCCCGTGAAGTGAACAGTGTGGAA	720			
QY	2358	TGGGCACTTCAACATTGACCAACCAAGCTCAGATTAAGTTCACTCTACAAAGTGTGAGC	2417			
DB	721	TGGGCACTTCAACATTGACCAACCAAGCTCAGATTAAGTTCACTCTACAAAGTGTGAGC	780			
QY	2418	CATGTGTAGAGCTGGCGGGCTGTGCTCAAGGCTGACCAAGACTTGCATGTGGCTGTG	2477			
DB	781	CATGTGTAGAGCTGGCGGGCTGTGCTCAAGGCTGACCAAGACTTGCATGTGGCTGTG	840			
QY	2478	CCAGGGCCAGGCCAGTGCACCTCTGCGCAGCACTGCCCTGCCAGAGAGCCAGTGGCT	2537			
DB	841	CCAGGGCCAGGCCAGTGCACCTCTGCGCAGCACTGCCCTGCCAGAGAGCCAGTGGCT	900			
QY	2538	GGAGGTGCTGTGTGCAAAAGCAAGTGCACAAACCCCGCATCAAGAGATTAATCCCGT	2597			
DB	901	GGAGGTGCTGTGTGCAAAAGCAAGTGCACAAACCCCGCATCAAGAGATTAATCCCGT	960			
QY	2598	GACAGAGCCCCCGGAAAGGGGACCAAGGTCACTATCCAGAGGAGAACTTGGGCTTGA	2657			


```

Db      961 GACAGGCCCCGGGAGAGGGGACCAAGGTCACTATCGAGGGGAGAACTGGGGCTGGA 1020
Qy      2658 ATTTTCGACATGCGCTTCCCATGTCAAGTTGCTGGGCGTGAAGTGCAGCCCTTTAGTGA 2717
Db      1021 ATTTCGGACATTCGCTCCCAATGCAAGTTGCTGGGCGTGAAGTGCAGCCCTTTAGTGA 1080
Qy      2718 TGTTCATCCCTGAGAACAGATCGTGTGAGATGGGGAGGCCAAGCCCAAGCCAGCA 2777
Db      1081 TGTTCATCCCTGAGAACAGATCGTGTGAGATGGGGAGGCCAAGCCCAAGCCAGCA 1140
Qy      2778 TGCAGGCTTCGTGAGATCTGCGTGGCTGTGTGTGCGCTGAATTCATGGCCCGGTCTC 2837
Db      1141 TGCAGGCTTCGTGAGATCTGCGTGGCTGTGTGTGCGCTGAATTCATGGCCCGGTCTC 1200
Qy      2838 ACAGCTCTATTAATTGACATGACATGACATCTCTCAGATCTGAAGCCCAAGCCGGGGCCCAT 2897
Db      1201 ACAGCTCTATTAATTGACATGACATGACATCTCTCAGATCTGAAGCCCAAGCCGGGGCCCAT 1260
Qy      2898 GTCCGAGGAGACCCAAAGTGAACATCAAGGACCAACCTGAATCCGGAAAGCAACGTGGT 2957
Db      1261 GTCCGAGGAGACCCAAAGTGAACATCAAGGACCAACCTGAATCCGGAAAGCAACGTGGT 1320
Qy      2958 GGTGATGTTTGGAAAGCAAGCCCTGTCTCTTCACAGGCGATCTGCATCTCAATTGCTG 3017
Db      1321 GGTGATGTTTGGAAAGCAAGCCCTGTCTCTTCACAGGCGATCTGCATCTCAATTGCTG 1380
Qy      3018 CAACACCAATCTCAGATGAGTGTCTAGATGAGTGTCTGAGTGTGAGTGTGACAGGAC 3077
Db      1381 CAACACCAATCTCAGATGAGTGTCTAGATGAGTGTCTGAGTGTGAGTGTGACAGGAC 1440
Qy      3078 CAAGATCAACAGGACCTGTCTTCAATGATGAGAAACCCCAACATCTGCGGATTTGA 3137
Db      1441 CAAGATCAACAGGACCTGTCTTCAATGATGAGAAACCCCAACATCTGCGGATTTGA 1500
Qy      3138 GCCAGATGAGACATTTGACGTGAAACACACCAATCGCCATGAGGGGACCCACCTGGA 3197
Db      1501 GCCAGATGAGACATTTGACGTGAAACACACCAATCGCCATGAGGGGACCCACCTGGA 1560
Qy      3198 CCTCATTAAGAAACCCCAAGTCCGTGCCAAGATGAGAGGAGAGACATCAATATCTG 3257
Db      1561 CCTCATTAAGAAACCCCAAGTCCGTGCCAAGATGAGAGGAGAGACATCAATATCTG 1620
Qy      3258 TGAGGTTCTGAAGCGTCACTGAGATGACCTGTGACGCGCCCTCGTCTGGGTCTCTGA 3317
Db      1621 TGAGGTTCTGAAGCGTCACTGAGATGACCTGTGACGCGCCCTCGTCTGGGTCTCTGA 1680
Qy      3318 CCACCAAGTGAACCTGACCGGAGAGGCGCGAGAGTTTGGCTTCACTCTGAGCAACGTCCA 3377
Db      1681 CCACCAAGTGAACCTGACCGGAGAGGCGCGAGAGTTTGGCTTCACTCTGAGCAACGTCCA 1740
Qy      3378 GTCCCTGCTCATCTCAACAGACCAACATCTCACTATCCCAACCGGATTTGAGGC 3437
Db      1741 GTCCCTGCTCATCTCAACAGACCAACATCTCACTATCCCAACCGGATTTGAGGC 1800
Qy      3438 CTTTGGTCCCTGAGAAATCTTGAAGCTCAAGCCTTGGACGCGCCATCATCTTAAAGGGCAA 3497
Db      1801 CTTTGGTCCCTGAGAAATCTTGAAGCTCAAGCCTTGGACGCGCCATCATCTTAAAGGGCAA 1860
Qy      3498 GAACTGATCCCGCTGTGGGCGTGGGGGACAGTGAAGTGAACATCACTGTGCGTGG 3557
Db      1861 GAACTGATCCCGCTGTGGGCGTGGGGGACAGTGAAGTGAACATCACTGTGCGTGG 1920
Qy      3558 GGAGAGACCGTGACACCGTGACCGTGTCAGATGTCCAGCTGCTTCCGAGTCCCAACCT 3617
Db      1921 GGAGAGACCGTGACACCGTGACCGTGTCAGATGTCCAGCTGCTTCCGAGTCCCAACCT 1980
Qy      3618 CATCGGACAGGCAAAAGTATGCGCGTGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGT 3677
Db      1981 CATCGGACAGGCAAAAGTATGCGCGTGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2040
Qy      3678 GTACATTTGCCCGGACAGGCGCGCTCAGCTGCGCCATCGTCAAGCATGCAAGTGGCTGG 3737

```

```

Db      2041 GTACATTTCCCGGACAGCCCGCTCAGCTGCGCCCATCGTCAAGCATGCAAGTGGCTGG 2100
Qy      3738 CGGCTCCCTCATCATTTTCATGCTGAGCGGTGCTCATGTGCTATTAACCAAGTCCCGGGA 3797
Db      2101 CGGCTCCCTCATCATTTTCATGCTGAGCGGTGCTCATGTGCTATTAACCAAGTCCCGGGA 2160
Qy      3798 AAGTGAACCTCAGCTGAAGCGGCTCAGATGAGATGAGCAACCTGAGTCCCTGTGGC 3857
Db      2161 AAGTGAACCTCAGCTGAAGCGGCTCAGATGAGATGAGCAACCTGAGTCCCTGTGGC 2220
Qy      3858 CCTGAGTGCAGAGAG 3874
Db      2221 CCTGAGTGCAGAGAG 2237

RESULT 4
BC028744
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

1 (bases 1 to 2015)
Strausberg,R.D., Fellnagold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Datchenko,L., Martusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Rana,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McMan,P.O.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Mortley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hilyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Sanchez,V., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
Boutard,G.G., Blakeley,R.W., Touchman,J.W., Green,S.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,J., Small,J.D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16893-16903 (2002)
22388257
12477932
2 (bases 1 to 2015)
Strausberg,R.
Direct Submission
Submitted (29-APR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.ncl.nih.gov
On Aug 25, 2003 this sequence version replaced gi:22382104.
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Miklos Palcovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiroki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.W.A.G.E. Consortium (LNL)
DNA sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadan@systemsbio.org
Amp Madan, Jessica Fahy, Erin Helton, Mark Kettelman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whitting
Clone distribution: MGC clone distribution information can be found

```

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 47 Row: B Column: 11.
 Location/Qualifiers
 1. 2015
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:34033 IMAGE:4828616"
 /tissue_type="Testis"
 /clone_id="NH1 MGC_97"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

gene
 1. 2015
 /gene="DKFZp434G0625"
 /db_xref="LOCUSID:91584"
 378. 1946
 /product="hypothetical protein DKFZp434G0625"
 /protein_id="AAH28744.1"
 /db_xref="GI:22382105"
 /db_xref="LOCUSID:91584"
 /translation="MKAMPNWTCLSLHLMVGGSSYLLTRQAPLRSKQSFYMR
 GEPAEGFNHLYDERTGHTLYGAVNRIYKLSYDLKLVTHETSPEDNKKCPRIYV
 TCNEPLITNNVNMKLLIDYENRLLACGSLYOGICLRLEDFLQSGPYRKHETL
 SGVNSGSPVGVIVSYNLDKLFIAVDGKPEPPTLSRKLTKNSPADSMFAVY
 HDEFVAMIKIPSDIPTIIPDPIYVYGFSSGFYFLIQEYVSPGSGTKEDY
 TSKLVRLKEDTAFNSYVEYVIGCSGVEYRLLQAAVYSKAAVETLGVYDDLL
 LPTVSKGQRKMKSLDESLALCTFIKKQINRKHRLKSGYEGEGLDAMVKKVDP
 CSKSLITIDNFGCLDMNAPLGVSDMVRGIPVTERPDVMTSVIAYVNRSLAFPT
 KSGKLSFSTGPGGTTQEMIGVAGDPFANAIASQPMCLVTLQSSSKAISDRVQ
 PLLCCFLNVPGSS"

misc_feature
 528. 1793
 /note="Sema; Region: semaphorin domain"
 /db_xref="CDD:smart00630"

ORIGIN

Query Match 24.1%; Score 1371; DB 9; Length 2015;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1371; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGAAGCGATGCGCTGGAAGTGAAGCTGCTCTCTCCACCTCTCATGTGTGGCATG 60
DB 378 ATGAAGCGATGCGCTGGAAGTGAAGCTGCTCTCTCCACCTCTCATGTGTGGCATG 437
QY 61 GGGTCTCTCACTTTGCTCAACCCGCGAGCGAGCCCGCTGTCCGAGAGCGAGGTCATT 120
DB 438 GGGTCTCTCACTTTGCTCAACCCGCGAGCGAGCCCGCTGTCCGAGAGCGAGGTCATT 497
QY 121 GTCAATTCGAGAGAGAGCGCGCGAGGGTTTCATCACTGCTGTGTGATGAGAGACA 180
DB 498 GTCAATTCGAGAGAGAGCGCGCGAGGGTTTCATCACTGCTGTGTGATGAGAGACA 557
QY 498 GTCAATTCGAGAGAGAGCGCGCGAGGGTTTCATCACTGCTGTGTGATGAGAGACA 557
DB 181 GAGACATTTACTTGGGGGCGCTCAATGCAATTTTCAAGGCTCTCCAGCGACTGAAGGTC 240
QY 558 GAGACATTTACTTGGGGGCGCTCAATGCAATTTTCAAGGCTCTCCAGCGACTGAAGGTC 617
DB 241 TTGTGAGCGATGAGAGAGAGCGCGAGAGAGCAACCCCAAGGTTCACCAACCCCGCATC 300
QY 618 TTGTGAGCGATGAGAGAGAGCGCGAGAGAGCAACCCCAAGGTTCACCAACCCCGCATC 677
DB 301 GTCCAGACCTGCAATGAGCGCGCTGACACCAACCAATGTCAACCAAGATGCTCTCAT 360
DB 678 GTCCAGACCTGCAATGAGCGCGCTGACACCAACCAATGTCAACCAAGATGCTCTCAT 737
QY 361 GACTCAAGAGAGAGAGCGCTGATTTCCCTGTGGAGCGCTGTACCAAGGATGTGCAAGCTG 420
DB 738 GACTCAAGAGAGAGAGCGCTGATTTCCCTGTGGAGCGCTGTGTACCAAGGATGTGCAAGCTG 797
QY 421 CTGAGGCTGAGAGAGCGCTTTCAGGCTGGGGAGGCTTATCATTAAGAAAGAGCACTATCTG 480
DB 798 CTGAGGCTGAGAGAGCGCTTTCAGGCTGGGGAGGCTTATCATTAAGAAAGAGCACTATCTG 857
QY 481 TAGGCTGACAGAGAGCGGCTCACTTTTGGAGTGAATGCTCTCTACAGCAACCTGGAT 540

```

```

DB 858 TGAAGTGTCAACGAGAGCGGCTCACTGCTTTTGGATGATCTCTCTCAAGCAACCTGGAT 917
QY 541 GACAAGCTGTTCATTTGCGACGCGAGTGAATGGGAAGCCCGAGATATTTTCCCAACCATCTCC 600
DB 918 GACAAGCTGTTCATTTGCGACGCGAGTGAATGGGAAGCCCGAGATATTTTCCCAACCATCTCC 977
QY 601 AGCCGGAACCTGACCAAGCAACTGTGAGCGCGATGAGTGTGGGTACGTTCTTCATGAT 660
DB 978 AGCCGGAACCTGACCAAGCAACTGTGAGCGCGATGAGTGTGGGTACGTTCTTCATGAT 1037
QY 661 GAGTTCGTGGCTGATGATTAAGATCCCTTGGACACTTTCACATCATCTCTGACTTT 720
DB 1038 GAGTTCGTGGCTGATGATTAAGATCCCTTGGACACTTTCACATCATCTCTGACTTT 1097
QY 721 GATATCTATCTATGCTCTATGCTTTAGAGATGGCACTTTGTCTACTTTTGAACCTCCAA 780
DB 1098 GATATCTATCTATGCTCTATGCTTTAGAGATGGCACTTTGTCTACTTTTGAACCTCCAA 1157
QY 781 CCTGAGATGATGTCTCCACAGGCTCCACCAAGAGAGAGAGTGTATACATCCAACTC 840
DB 1158 CCGTGAAGTGTGTCTCCACAGGCTCCACCAAGAGAGAGAGTGTATACATCCAACTC 1217
QY 841 GTGAGGCTTTGCAAGAGAGACACAGCTTCACTCCATGATGAGTGGCCATTTGGCTGT 900
DB 1218 GTGAGGCTTTGCAAGAGAGACACAGCTTCACTCCATGATGAGTGGCCATTTGGCTGT 1277
QY 901 GAGCGCACTGGGAGTGAAGTACCGGCTGTGAGGCTGCTTCACTGTCCAAAGCGGAGGCC 960
DB 1278 GAGCGCACTGGGAGTGAAGTACCGGCTGTGAGGCTGCTTCACTGTCCAAAGCGGAGGCC 1337
QY 961 GTGCTTGGCAGAGACCTTTGAGAGTCCATTCAGATGATACCTGCTCTTCAACGCTCTCC 1020
DB 1338 GTGCTTGGCAGAGACCTTTGAGAGTCCATTCAGATGATACCTGCTCTTCAACGCTCTCC 1397
QY 1021 AAGGCGCAGAAAGCGGAAATGAAATCCCTGATGATGCGGCTGTGATCTTCATCTG 1080
DB 1398 AAGGCGCAGAAAGCGGAAATGAAATCCCTGATGATGCGGCTGTGATCTTCATCTG 1457
QY 1081 AAGCAGATTAATGACCGGATTAAGAGCGGCTGAGTCTTTACCGGGGAGAGGCAAG 1140
DB 1458 AAGCAGATTAATGACCGGATTAAGAGCGGCTGAGTCTTTACCGGGGAGAGGCAAG 1517
QY 1141 CTGACCTGGGCTGCTCAAGGTGAGAGACATCCCTGACAGTGGCTCTTAACAT 1200
DB 1518 CTGACCTGGGCTGCTCAAGGTGAGAGACATCCCTGACAGTGGCTCTTAACAT 1577
QY 1201 GACGATTAATCTGTGGCTTGAAGATGATCTCCCTGGAGTGTCCGACATGGTGGT 1260
DB 1578 GACGATTAATCTGTGGCTTGAAGATGATCTCCCTGGAGTGTCCGACATGGTGGT 1637
QY 1261 GGAATTCGCGCTTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
DB 1638 GGAATTCGCGCTTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1697
QY 1321 AAGAACCACTCTCTGAGCTTTGTGGGACCAAAAGTGGCAAGCTGAAGAG 1371
DB 1698 AAGAACCACTCTCTGAGCTTTGTGGGACCAAAAGTGGCAAGCTGAAGAG 1748

```

RESULT 5
 AK123428 3310 bp mRNA linear PRI 09-SEP-2003
 LOCUS AK123428
 DEFINITION Homo sapiens cDNA FLJ41434 f1s, clone BH192007616, moderately
 similar to P1exin 2.
 ACCESSION AK123428
 VERSION AK123428.1 GI:34528974
 KEYWORDS oligo capping; f1s (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1

AC009785/c 179145 bp DNA linear PRI 26-MAY-2002
 LOCUS AC009785
 DEFINITION Homo sapiens chromosome 7 clone RP11-198E23, complete sequence.
 ACCESSION AC009785
 VERSION AC009785.8 GI:21217401
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE AUTHORS
 1 Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
 Saepthimachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C.
 and Haugen, E.D.
 TITLE Direct Submission
 JOURNAL Unpublished
 2 (bases 1 to 179145)
 REFERENCE AUTHORS
 Bubb, K.L., Desmarais, C.L., Ramsey, S.A. and Hudley, R.M.
 TITLE Direct Submission
 JOURNAL Submitted (01-SEP-1999) Human Genome Center, University of
 Washington, Box 352145, Seattle, WA 98195, USA
 3 (bases 1 to 179145)
 REFERENCE AUTHORS
 Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Raymond, C.,
 Cleidenning, J., Ivey, R.G. and Haugen, E.D.
 TITLE Direct Submission
 JOURNAL Submitted (20-JUN-2001) Genome Center, University of Washington,
 Box 352145, Seattle, WA 98195, USA
 4 (bases 1 to 179145)
 REFERENCE AUTHORS
 Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
 Saepthimachak, C., Phelps, K.A., Raymond, C. and Haugen, E.D.
 TITLE Direct Submission
 JOURNAL Submitted (06-APR-2002) Genome Center, University of Washington,
 Box 352145, Seattle, WA 98195, USA
 5 (bases 1 to 179145)
 REFERENCE AUTHORS
 Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
 Saepthimachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C.
 and Haugen, E.D.
 TITLE Direct Submission
 JOURNAL Submitted (26-MAY-2002) Genome Center, University of Washington,
 Box 352145, Seattle, WA 98195, USA
 COMMENT On May 26, 2002 this sequence version replaced gi:20066281.
 Center: University of Washington Genome Center
 Center Code: UWGC
 Web site: <http://www.genome.washington.edu>
 Contact: uwgchgs@u.washington.edu
 Project Information
 Center project name: chr-7
 Center clone name: RP11-198E23 (djs380)
 Summary Statistics
 Sequencing vector: plasmid; X52328, 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 178850 bases at least Q40
 Consensus quality: 179053 bases at least Q30
 Consensus quality: 179141 bases at least Q20
 Insert size: 179145; sum-of-contigs
 Quality coverage: 11.8x in Q20 bases; sum-of-contigs

Overlapping Sequences:

5' : RP11-341C17 (UWGC:djs301) AC011625 73322-bp overlap
 3' : RP11-335B7 (UWGC:djs42) AC026239 38206-bp overlap

Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted:
 all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M3 subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

EcoRI				HindIII				BglII			
SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt
8696	8850	399	<800	2304	2280	8696	8850	399	<800	2304	2280
6	<800	6382	6347	2067	2069	6	<800	6382	6347	2067	2069
1907	1895	512	<800	5518	5637	1907	1895	512	<800	5518	5637
1854	1895	449	<800	6918	7002	1854	1895	449	<800	6918	7002
2849	2922	2322	2276	2029	2069	2849	2922	2322	2276	2029	2069
3722	3806	5951	5995	2378	2425	3722	3806	5951	5995	2378	2425
658	<800	7224	7291	2089	2069	658	<800	7224	7291	2089	2069
26398	26593	1005	1033	550	<800	26398	26593	1005	1033	550	<800
537	<800	1020	1033	4191	4187	537	<800	1020	1033	4191	4187
1542	1483	3445	3448	3839	3773	1542	1483	3445	3448	3839	3773
1103	1096	2131	2126	1647	1646	1103	1096	2131	2126	1647	1646
1928	2045	2237	2276	629	<800	1928	2045	2237	2276	629	<800
14029	13827	6625	6652	1932	1918	14029	13827	6625	6652	1932	1918
6707	6816	6349	6347	18266	18041	6707	6816	6349	6347	18266	18041
1884	1895	6669	6652	3099	3165	1884	1895	6669	6652	3099	3165
6048	6029	1943	1916	2900	2928	6048	6029	1943	1916	2900	2928
2150	2209	6376	6347	111	<800	2150	2209	6376	6347	111	<800
13944	13827	5398	5641	2425	2425	13944	13827	5398	5641	2425	2425
4275	4287	599	<800	3767	3773	4275	4287	599	<800	3767	3773
2057	2209	4055	4127	1506	1485	2057	2209	4055	4127	1506	1485
2293	2209	299	<800	3639	3773	2293	2209	299	<800	3639	3773
6448	6393	20065	20079	5635	5637	6448	6393	20065	20079	5635	5637
9916	9943	6964	6903	6894	7002	9916	9943	6964	6903	6894	7002
7673	7645	5766	5641	11819	11807	7673	7645	5766	5641	11819	11807

115	<800	4281	4127	837	821
1490	1483	7844	7901	2058	2069
1071	1096	1032	1033	3280	3312
18341	18612	1591	1546	3780	3773
7618	7645	955	954	623	<800
1733	1700	330	<800	202	<800
1478	1483	1927	1916	2950	2928
7616	7645	6261	6347	10775	10621
561	<800	445	<800	12114	11807
19080	18612	3869	3877	9445	9461
114	<800	4148	4127	1694	1646
7	<800	3828	3877	3745	3773
		4001	4127	674	<800
		1561	1546	9893	9817
		3188	3192	3225	3165
		4960	4943	998	1011
		9109	8931	1318	1283
		929	954	94	<800
		7844	7901	796	821
		535	<800	111	<800
		5397	5408	1029	1011
		1905	1916	1491	1485
		113	<800	20564	20918
		5596	5641		
		750	735		
		1264	1248		

FEATURES
Source
Location/Qualifiers
1. 179145
/organism="Homo sapiens"

Query Match
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGAAAGCCATGCCCTGGAACCTGACCTGCTTCTTCCCACTCTCTCAATGATGGGATG 60
DB	61796	ATGAAAGCCATGCCCTGGAACCTGACCTGCTTCTTCCCACTCTCTCAATGATGGGATG 61737
QY	61	GGCTCTCCACTTTGCTCAACCGGAGGACCGCCGCTGTCCTCCAGAGACAGGCTCATTT 120
DB	61736	GGCTCTCCACTTTGCTCAACCGGAGGACCGCCGCTGTCCTCCAGAGACAGGCTCATTT 61677
QY	121	GTCACATTCCAG 180
DB	61676	GTCACATTCCAG 61617
QY	181	GGAACATTACTGGGGGCGGTGAATGATTTAACAAGCTCTCCAGAGACCTGAAGGTC 240

DB	61616	GGAACATTACTGGGGGCGGTGAATGATTTAACAAGCTCTCCAGAGACCTGAAGGTC 61557
QY	241	TTGGTGAAGCATGAGACAGAGGCGCGGACGAGAGACACCCCAAGTGTTCACCCCGCATC 300
DB	61556	TTGGTGAAGCATGAGACAGAGGCGCGGACGAGAGACACCCCAAGTGTTCACCCCGCATC 61497
QY	301	GTCAGACCTGCAATGAGCCCTGACACACACCAATGTCAACAATGCTCTCTATA 360
DB	61496	GTCAGACCTGCAATGAGCCCTGACACACACCAATGTCAACAATGCTCTCTATA 61437
QY	361	GACTACAG 420
DB	61436	GACTACAG 61377
QY	421	CTGAGGCTGAG 480
DB	61376	CTGAGGCTGAG 61317
QY	481	TCAGGTGTCAACGAGAGGCGCTCAGTCTTTGAGTATGCTCTCTCAAGCAACCTGAT 540
DB	61316	TCAGGTGTCAACGAGAGGCGCTCAGTCTTTGAGTATGCTCTCTCAAGCAACCTGAT 61257
QY	541	GACAGGCTGTCATTGCCAGGCGGAGTGAATGGAAGCCGAGTATTTTCCACCATCTCC 600
DB	61256	GACAGGCTGTCATTGCCAGGCGGAGTGAATGGAAGCCGAGTATTTTCCACCATCTCC 61197
QY	601	AGCCGGAACCTGACCAAGAACTCTGAGCGGAGTGGCATGTTGCGGTACGTTCCATGAT 660
DB	61196	AGCCGGAACCTGACCAAGAACTCTGAGCGGAGTGGCATGTTGCGGTACGTTCCATGAT 61137
QY	661	GAGTTCGGGCTGCAATGATTAAGTCCCTCGAGACCTTCACCATCATCCGATCTT 720
DB	61136	GAGTTCGGGCTGCAATGATTAAGTCCCTCGAGACCTTCACCATCATCCGATCTT 61077
QY	721	GATATCTACTATGCTATGCTTATGCTTTAGCAGTGGCAACTTGTCTATTTTACCCTCAA 780
DB	61076	GATATCTACTATGCTATGCTTATGCTTTAGCAGTGGCAACTTGTCTATTTTACCCTCAA 61017
QY	781	CCTGAGATGATGCTCTCCACAGGCTCCACACAGAGAGAGAGAGAGAGAGAGAGAGAG 840
DB	61016	CCTGAGATGATGCTCTCCACAGGCTCCACACAGAGAGAGAGAGAGAGAGAGAGAGAG 60957
QY	841	GTAAGGCTTTGCAAG 900
DB	60956	GTAAGGCTTTGCAAG 60897
QY	901	GAGCGAGTGGGCTGAGTACCGCTGCTGCAAGGCTGCTTACCTGTCCAAAGCGGGGCC 960
DB	60896	GAGCGAGTGGGCTGAGTACCGCTGCTGCAAGGCTGCTTACCTGTCCAAAGCGGGGCC 60837
QY	961	GTCCTTGGAG 1020
DB	60836	GTCCTTGGAG 60777
QY	1021	AGGCGCCAG 1080
DB	60776	AGGCGCCAG 60717
QY	1081	AAGCAGATTAATGACCGGATTAAGAGAGGCGCTGAGTCTTGTTCAGGGGAGAGGACAG 1140
DB	60716	AAGCAGATTAATGACCGGATTAAGAGAGGCGCTGAGTCTTGTTCAGGGGAGAGGACAG 60657
QY	1141	CTGAGACCTGGCTGCTCAAGGTGAAGACATCCCTTCAGAGAGTGC 1188
DB	60656	CTGAGACCTGGCTGCTCAAGGTGAAGACATCCCTTCAGAGAGTGC 60609

RESULT 7
AX833678
LOCUS AX833678 3666 bp DNA linear PAT 15-DEC-2003
DEFINITION Sequence 802 from Patent EP1347046.
ACCESSION AX833678

VERSION AX833678.1 GI:39919813
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 Isogai, T., Sugiyama, T., Otsuki, T., Makamatsu, A., Sato, H., Ishii, S., Yamamoto, J., I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R., Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahara, K. and Masuh, Y.
 Full-length cDNA sequences
 Patent: EP 1347046-A 802 24-SEP-2003;
 Research Association for Biotechnology (JP)
 FEATURES
 source 1..3666
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Query Match 19.5%; Score 1108; DB 6; Length 3666;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1158; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 4527 TGCCACAGCCCGAGGAGTCCAGTAAGATCTCAACTGTGACACCATCACTCAGGTCAA 4586
 DB 1 TGCCACAGCCCGAGGAGTCCAGTAAGATCTCAACTGTGACACCATCACTCAGGTCAA 50
 QY 4587 GAGAAAGATTTCTGATGCCATCTTCAAGATGTGCTGCTCCACCGGCGCAAGCTGC 4646
 DB 61 GAGAAAGATTTCTGATGCCATCTTCAAGATGTGCTGCTCCACCGGCGCAAGCTGC 120
 QY 4647 AGATATGATCTGAGTGGCGACAGAGAGTGGGCGCAAGGATGATCTTGCAGGATGAAGA 4706
 DB 121 AGATATGATCTGAGTGGCGACAGAGAGTGGGCGCAAGGATGATCTTGCAGGATGAAGA 180
 QY 4707 CATACACCAAGATTGAGAAATGATTGAGAGCCATGTAACAACCTGCGCCATACCAAGGT 4766
 DB 181 CATACACCAAGATTGAGAAATGATTGAGAGCCATGTAACAACCTGCGCCATACCAAGGT 240
 QY 4767 GCCAGATGATTCCTGCTGAGTGGCATGATGTCTCAAGCAGAGTGAACGCTATACCGCACTGAA 4826
 DB 241 GCCAGATGATTCCTGCTGAGTGGCATGATGTCTCAAGCAGAGTGAACGCTATACCGCACTGAA 300
 QY 4827 CAATCTCCACGCTCTCCAGAGCTTCAAGCAAGTAAATATGAAACATGATCCGGTACACGGG 4886
 DB 301 CAATCTCCACGCTCTCCAGAGCTTCAAGCAAGTAAATATGAAACATGATCCGGTACACGGG 360
 QY 4887 CAGCCCGCAGACGCTCCGCTCAAGAGACACCTATGATCACTCTGACCTGAGAGTGGAGT 4946
 DB 361 CAGCCCGCAGACGCTCCGCTCAAGAGACACCTATGATCACTCTGACCTGAGAGTGGAGT 420
 QY 4947 CAAGATGTGCACTTAACTGTAAGAAACAACAGAGACCGAGACCAAGAGAGGCGGAGCGGG 5006
 DB 421 CAAGATGTGCACTTAACTGTAAGAAACAACAGAGACCGAGACCAAGAGAGGCGGAGCGGG 480
 QY 5007 GAGCAAGATGTGTCTGTAATCTTCACTGACCCGACTCTTGCCACTTAAGGCGACACTGCA 5066
 DB 481 GAGCAAGATGTGTCTGTAATCTTCACTGACCCGACTCTTGCCACTTAAGGCGACACTGCA 540
 QY 5067 GAAAGTTGTGATGATGATCTTCTTGAACATCTTCAAGACCGGACACCGGTGCTTGCCCT 5126
 DB 541 GAAAGTTGTGATGATGATCTTCTTGAACATCTTCAAGACCGGACACCGGTGCTTGCCCT 600
 QY 5127 GCCCTTGCGCCATCAAGTACATGTTTGAATCTTCTGATGAGAGGCTGATTAACATGGCAT 5186
 DB 601 GCCCTTGCGCCATCAAGTACATGTTTGAATCTTCTGATGAGAGGCTGATTAACATGGCAT 660
 QY 5187 TCATGACCCGACGTCGCGCATACCTGTAAGAGCAATTCCTGCGCCCTGAGAGTTTGGGT 5246
 DB 661 TCATGACCCGACGTCGCGCATACCTGTAAGAGCAATTCCTGCGCCCTGAGAGTTTGGGT 720

QY 5247 CAACATGATCAAGAACCCGACGATTTGTGTTGACATCCATAGAACAGCATCAAGACGC 5306
 DB 721 CAACATGATCAAGAACCCGACGATTTGTGTTGACATCCATAGAACAGCATCAAGACGC 780
 QY 5307 CTGCTCTCTGTGTGTGCTGACACTTCAATGAGACTCTTCTCTCCAGCTCAAGACCGGCT 5366
 DB 781 CTGCTCTCTGTGTGTGCTGACACTTCAATGAGACTCTTCTCTCCAGCTCAAGACCGGCT 840
 QY 5367 GGGCAAGAGCTCGCCCTCCAAAGAGCTGTGTATGATCCAGAGACATCCCGAGTACAGAA 5426
 DB 841 GGGCAAGAGCTCGCCCTCCAAAGAGCTGTGTATGATCCAGAGACATCCCGAGTACAGAA 900
 QY 5427 TTGGGTGAGAGGATTTACTCAACATAGAGAGATGCCAGCATCAAGCCACAGACAT 5486
 DB 901 TTGGGTGAGAGGATTTACTCAACATAGAGAGATGCCAGCATCAAGCCACAGACAT 960
 QY 5487 GAAGCATACCTGCTGAGAGAGTCCGAGATGACATGATGATGATCAACACCATGAGTGC 5546
 DB 961 GAAGCATACCTGCTGAGAGAGTCCGAGATGACATGATGATGATCAACACCATGAGTGC 1020
 QY 5547 ACTCTCAGAGATCTTCTCTATATGTGGCAATACAGCGAGAGATCTTGGACCTTGA 5606
 DB 1021 ACTCTCAGAGATCTTCTCTATATGTGGCAATACAGCGAGAGATCTTGGACCTTGA 1080
 QY 5607 CCAAGATACCAAGTGTGGAGAGCAAGAACGCTTACAACTGAACAGCATTAACCT 5666
 DB 1081 CCAAGATACCAAGTGTGGAGAGCAAGAACGCTTACAACTGAACAGCATTAACCT 1140
 QY 5667 CATGAGCTTACAGAGCTGA 5685
 DB 1141 CATGAGCTTACAGAGCTGA 1159
 RESULT 8
 AK095606 3666 bp mRNA linear PRI 15-JUN-2002
 LOCUS AK095606
 DEFINITION Homo sapiens cDNA FLJ38287 f1s, clone FCBBP3008362, moderately similar to PLEXIN 4 PRECURSOR.
 ACCESSION AK095606
 VERSION AK095606.1 GI:21754899
 KEYWORDS Oligo capping, fls (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 Suzuki, O., Sasaki, N., Aotsuma, S., Shoji, T., Ichihara, T., Shiohara, N., Matsumoto, K., Hirano, M., Sano, S., Nemura, R., Yoshikawa, Y., Matsumura, Y., Moriya, S., Chiba, E., Momiyama, H., Onogawa, S., Kaeriyama, S., Satoh, N., Matsunawa, H., Takahashi, E., Katoka, R., Kuga, N., Kuroda, A., Satoh, I., Kamata, K., Takami, S., Terashima, Y., Watanabe, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Ota, T., Makamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai, Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagaetsuma, M., Murakawa, K., Kanehori, K., Takahashi, Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuh, Y., Nagai, K. and Isogai, T.
 NEDO human cDNA sequencing project
 TITLE NEDO human cDNA sequencing project
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 3666)
 AUTHORS Isogai, T. and Yamamoto, J.
 DIRECT SUBMISSION
 TITLE Submitted (04-JUN-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
 COMMENT (E-mail:genominfo@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and

RAB; annotation: HRI and RAB.

FEATURES	Location/Qualifiers
source	1. .3666

Source

```

/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="FCBSP3100862"
/risseq_type="Drain"
/clone_id="FCBSP3"
/dev_stage="fetal"
/note="Cloning vector: pME18SF3"
125.1159

```

ORIGIN

Query Match	19.5%	Score 1106;	DB 9;	Length 3666;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 1158;	Conservative	0;	Mismatches 1;	Indels 0;
				Gaps 0;

QY	4527	TGCGAACACCCCGGAGGTCCCGAAGAAAGATCTCAACCTGGACACCATCACTCAGGTCA	4588
Db	1	TGCGAACACCCCGGAGGTCCCGAAGAAAGATCTCAACCTGGACACCATCACTCAGGTCA	60
QY	4587	GGAGAAAGATTCTGATATGCCATCTTCAGAAATATGTGCTTGTGCCACCGGCGAAAGCTGC	4648
Db	61	GGAGAAAGATTCTGATATGCCATCTTCAGAAATATGTGCTTGTGCCACCGGCGAAAGCTGC	120
QY	4647	AGATATGATCTGAGATGCGCGACAGAGAAATGTGGGCAAGAGATATCTTCAGAGATGAAGA	4708
Db	121	AGATATGATCTGAGATGCGCGACAGAGAAATGTGGGCAAGAGATATCTTCAGAGATGAAGA	180
QY	4707	CATCACACCAACAGATTGAGAAATGATGGGAACGCACTGAACAACCTGGGCCACATACAGGT	4766
Db	181	CATCACACCAACAGATTGAGAAATGATGGGAACGCACTGAACAACCTGGGCCACATACAGGT	240
QY	4767	GCCAGATGATTCGGTGTGGCATTAATGATCTCCAGCAGGTGAACAGCTATPACGAGGTGA	4828
Db	241	GCCAGATGATTCGGTGTGGCATTAATGATCTCCAGCAGGTGAACAGCTATPACGAGGTGA	300
QY	4827	CAACTCCACCGGTCCAGGACCTCAGCAAGTAATATGAAAACATGATCCGGTACACGGG	4888
Db	301	CAACTCCACCGGTCCAGGACCTCAGCAAGTAATATGAAAACATGATCCGGTACACGGG	360
QY	4887	CAGCCCGACAGCTCTCGCTCAAGCAACCTATGATCATCTCTGACCTGGACAGTGAAT	4944
Db	361	CAGCCCGACAGCTCTCGCTCAAGCAACCTATGATCATCTCTGACCTGGACAGTGAAT	420
QY	4947	CAAGATATGTGACCTTAGTGAAGAACACAGACACCGAGACCGAAGAGAGGGGACCCGGG	5006
Db	421	CAAGATATGTGACCTTAGTGAAGAACACAGACACCGAGACCGAAGAGAGGGGACCCGGG	480
QY	5007	GACCAAGATGTGTCTTGAAATCTACTCTAACCCGACTCTGGGCCATPAAAGGGCACTGCA	5066
Db	481	GACCAAGATGTGTCTTGAAATCTACTCTAACCCGACTCTGGGCCATPAAAGGGCACTGCA	540
QY	5067	GAAATTTGTGATGACCTCTTTGAGAACATCTTGACACGGCACACCTGTGGCTTGTGCTT	5128
Db	541	GAAATTTGTGATGACCTCTTTGAGAACATCTTGACACGGCACACCTGTGGCTTGTGCTT	600
QY	5127	GCCCTGGCCATCAATPACATGTTGACCTTCCTGGATATAGCAGGCTATPAAACATGGCAT	5188
Db	601	GCCCTGGCCATCAATPACATGTTGACCTTCCTGGATATAGCAGGCTATPAAACATGGCAT	660

QY	518	TCATGACCCGACAGTCCGACATACCTCGAGAAAGACCAATGCTCCGCTGAGGTTTGGGT	5246
Db	661	TCATGACCCGACAGTCCGACATACCTCGAGAAAGACCAATGCTCCGCTGAGGTTTGGGT	720
QY	5247	CAACATGATCAAGAACCCCGAGTTTGTTGAATCCATCAAGAACAGCATCAACAAGC	5306
Db	721	CAACATGATCAAGAACCCCGAGTTTGTTGATCATCCATMAAGAACGATCAACAAGC	780
QY	5307	CTGCGCTCTCTGTGGTGGCTCAGACCTTCATGAGACTCTGTGCTCCACGTCACAGACCGAGT	5366
Db	781	CTGCGCTCTCTGTGGTGGCTCAGACCTTCATGAGACTCTGTGCTCCACGTCACAGACCGAGT	840
QY	5367	GGGCAAGAGACTGCGCCCTCCACAAGCTGTGTATGCCAAAGACATCCCAAGCTTAAAGAA	5426
Db	841	GGGCAAGAGACTGCGCCCTCCACAAGCTGTGTATGCCAAAGACATCCCAAGCTTAAAGAA	900
QY	5427	TTGGGTGGAGAGGATTTACTCAGACATAGGAGAAAGATCCAGCCATCAGCAGCAAGAAT	5486
Db	901	TTGGGTGGAGAGGATTTACTCAGACATAGGAGAAAGATCCAGCCATCAGCAGCAAGAAT	960
QY	5487	GAAGCATACTGGCTGAGCAGAGTCCCGATGCACTGAATGAGTTCAACACCATGAGTGC	5546
Db	961	GAAGCATACTGGCTGAGCAGAGTCCCGATGCACTGAATGAGTTCAACACCATGAGTGC	1020
QY	5547	ACTTTCAGAAATCTTTCCTCCATGTGGGCAATAACAGGAGAGAGATCCTTGAGACTCTGGA	5606
Db	1021	ACTTTCAGAAATCTTTCCTCATGTGGGCAATAACAGGAGAGAGATCCTTGAGACTCTGGA	1080
QY	5607	CCAAGATGACCAAGTGTGGAGAGCAAAATCGGCTTACAACTAGAACAAAGTCATTAACCT	5666
Db	1081	CCAAGATGACCAAGTGTGGAGAGCAAAATCGGCTTACAACTAGAACAAAGTCATTAACCT	1140
QY	5667	CATGAGCTTAGACAGCTGA	5685
Db	1141	CATGAGCTTAGACAGCTGA	1159

	RESULT 9	
AC011625/c	LOCUS	AC011625 188090 bp DNA linear PRI 08-DEC-1999
	DEFINITION	Homo sapiens clone RP11-341C17 from 7q21, complete sequence.
	ACCESSION	AC011625
	VERSION	AC011625.2 GI:6539285
	KEYWORDS	HTG.
SOURCE	ORGANISM	Homo sapiens (human)
		Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 188090) Kaul,R.K., Yu,U., Wong,G.K.-S., Magness,C.L., Green,E.D., Green,F. and Olson,M.V. Large-scale MCD Mapping and Sequencing of Human Chromosome 7 Unpublished 2 (bases 1 to 188090) Bubb,K.L., Desmarais,C.L., Ramsey,S.A. and Hubley,R.M. Direct Submission Submitted (08-OCT-1999) Human Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA 3 (bases 1 to 188090) Kaul,R.K. and Desmarais,C.L. Direct Submission Submitted (08-DEC-1999) Human Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA On Dec 8, 1999 this sequence version replaced gi:6016750.
COMMENT		

----- Genome Center:
University of Washington Genome Center
Center code: UWGC
Web site: <http://genome.washington.edu>
Contact: uwgchelp@u.washington.edu
----- Project Information
Center project name: chr-7
Center clone name: djs301 (RP11-141C17)

Summary Statistics

Sequencing vector: M13, 100% of reads
Chemistry: Dye-Primer Bodipy, 93% of reads Chemistry:
Dye-Terminator Big Dye, 7% of reads
Assembly program: Phrap, version 0.990319
Insert size: 188 094; sum-of-contigs
Quality coverage: 7.78x in Q20 bases; sum-of-contigs

Overlapping Sequences:

5 : UMGc:dj8302
3 : UMGc:dj8380

Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the Genbank flat file format but are available as part of this entry's ASN.1 file.

Double stranded (DS) coverage: 70.4%
DS or two chemistry coverage: 100.0%
Single stranded regions: 0

Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-600 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

FP Seq FP Seq FP Seq
Bg111 Bc0r1 Ns11

2095.00	2061.00	8947.00	8687.00	729.00	732.00
12518.00	12347.00	844.00	812.00	1463.00	1480.00
5082.00	5007.00	4657.00	4593.00	5528.00	5438.00
2694.00	2672.00	3637.00	3613.00	3383.00	3371.00
5175.00	5083.00	1152.00	1147.00	1240.00	1220.00
1094.00	1093.00	11517.00	11411.00	4875.00	4844.00
2231.00	2189.00	16729.00	16537.00	4194.00	4094.00
16006.00	15888.00	13550.00	13383.00	2402.00	2353.00
6982.00	6668.00	5001.00	4900.00	1463.00	1439.00
8822.00	8629.00	6353.00	6107.00	2072.00	2056.00
11945.00	11811.00	3101.00	3027.00	5301.00	5143.00
10740.00	10587.00	3850.00	3719.00	1463.00	1408.00
2095.00	2062.00	1925.00	1935.00	729.00	706.00
7981.00	7820.00	6169.00	6045.00	10419.00	10239.00
2095.00	2101.00	6353.00	6206.00	4194.00	4107.00
1178.00	1141.00	1743.00	1714.00	1128.00	1118.00
3623.00	3525.00	2067.00	2052.00	622.00	630.00

621.00	581.00	559.00	569.00	1045.00	1052.00
6428.00	6341.00	544.00	510.00	7407.00	7153.00
6146.00	6023.00	2364.00	2354.00	3498.00	3389.00
769.00	736.00	660.00	633.00	4506.00	4428.00
4722.00	4640.00	1013.00	972.00	2402.00	2339.00
1178.00	1180.00	11962.00	11788.00	13094.00	12940.00
7083.00	6917.00	5201.00	5141.00	4378.00	4366.00
2095.00	2022.00	5201.00	5091.00	2025.00	1988.00
2429.00	2372.00	1925.00	1902.00	4037.00	3911.00
2095.00	2083.00	1925.00	1853.00	2402.00	2326.00
546.00	544.00	2967.00	2843.00	2072.00	2069.00
4240.00	4182.00	3850.00	3715.00	1609.00	1583.00
3913.00	3829.00	660.00	652.00	2922.00	2845.00
1656.00	1641.00	29666.00	26384.00	5301.00	5252.00
621.00	623.00	544.00	531.00	1747.00	1719.00
1940.00	1927.00	1551.00	1536.00	729.00	697.00
18696.00	18261.00	1152.00	1097.00	2646.00	2580.00
3153.00	3085.00	1925.00	1922.00	2402.00	2335.00
2966.00	2898.00	14113.00	14024.00	10419.00	10108.00
2429.00	2419.00	6876.00	6704.00	7407.00	7317.00
3913.00	3763.00	1925.00	1869.00	8785.00	8640.00
1524.00	1500.00	6169.00	6048.00	7052.00	6870.00
3776.00	3636.00	2207.00	2144.00	3731.00	3659.00
5704.00	5638.00			4194.00	4085.00
6982.00	6868.00			622.00	581.00
				3881.00	3751.00
				1240.00	1221.00
				1319.00	1284.00
				5528.00	5416.00
				7052.00	6844.00
				3881.00	3888.00
				3458.00	3426.00
				14516.00	14342.00

FEATURES

SOURCE

Location/Qualifiers
1. 188090
/Organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="7"

```
/map="7q31"
/clone="RP11-341C17 or IMGC-dj301"
/cell_line="Male Blood"
/clone_id="RPC-II Human Male BAC Library"
444..750
repeat_region
/rpc_family="Alu"
891..963
/strand name="SMS2952"
/ncbi="Genbank Accession: G13043"
complement(2131..2409)
/rpc_family="Alu"
repeat_region
complement(8590..8831)
/rpc_family="Alu"
10545..10834
repeat_region
/rpc_family="Alu"
11525..11795
repeat_region
/rpc_family="Alu"
18103..18403
repeat_region
/rpc_family="Alu"
complement(27089..27369)
repeat_region

Query Match 19.4%; Score 1106; DB 9; Length 188090;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1156; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAAAGCCATGCCCCGGAACCTGACCTGCTCTCCCACTCTCTCATGTGTGGGATG 60
DB 176564 ATGAAAGCCATGCCCCGGAACCTGACCTGCTCTCTCCCACTCTCTCATGTGTGGGATG 176505
QY 61 GGCTCTCCCACTTGTGTCACCCCGGACCGACCCCGCTGTCCAGAGAGAGAGAGATTT 120
DB 176504 GGCTCTCCCACTTGTGTCACCCCGGACCGACCCCGCTGTCCAGAGAGAGAGATTT 176445
QY 121 GTCAATTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
DB 176444 GTCAATTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 176385
QY 181 GGAACATTTACTTGGGGGCGCTCAATCGGATTTTAAAGCTTTCACAGGACCTGAAAGTC 240
DB 176384 GGAACATTTACTTGGGGGCGCTCAATCGGATTTTAAAGCTTTCACAGGACCTGAAAGTC 176325
QY 241 TTGGTGAAGCATGAGACAGGGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
DB 176324 TTGGTGAAGCATGAGACAGGGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 176265
QY 301 GTCCAGACCTGCAATGAGCCCTTGAACCAACCAACATGTCAACAAAGATGCTCTCAT 360
DB 176264 GTCCAGACCTGCAATGAGCCCTTGAACCAACCAACATGTCAACAAAGATGCTCTCAT 176205
QY 361 GACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 176204 GACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 176145
QY 421 CTGAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 176144 CTGAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 176085
QY 481 TCAGGTGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
DB 176084 TCAGGTGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 176025
QY 541 GACAAAGCTGTTCAATGCGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
DB 176024 GACAAAGCTGTTCAATGCGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 175965
QY 601 AGCCGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
DB 175964 AGCCGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 175905
QY 661 GAGTTCGTGGCTCGATGATTAAGATCCCTTCCGACACCTTCAACATCATCCCTGACTTT 720
DB 175904 GAGTTCGTGGCTCGATGATTAAGATCCCTTCCGACACCTTCAACATCATCCCTGACTTT 175845
```

```
QY 721 GATATCTACTATGCTATGTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
DB 175844 GATATCTACTATGCTATGTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 175785
QY 781 CCTGAGATGATGCTCTCCACCGAGGCTCCACCGACAGAGAGAGAGAGAGAGAGAGAG 840
DB 175784 CCTGAGATGATGCTCTCCACCGAGGCTCCACCGACAGAGAGAGAGAGAGAGAGAGAG 175725
QY 841 GTGAGAGCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
DB 175724 GTGAGAGCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 175665
QY 901 GAGCGCAGTGGGGTGGAGAGTACCGCTGCTGCAAGGCTGCTTACCTGTTCCAAAGCGGGGCGC 960
DB 175664 GAGCGCAGTGGGGTGGAGAGTACCGCTGCTGCAAGGCTGCTTACCTGTTCCAAAGCGGGGCGC 175605
QY 961 GTGCTTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
DB 175604 GTGCTTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 175545
QY 1021 AAGGCGCAGAAAGCGGAAATGAAATCCCTGATGAGTGGCGCTGTGCACTTTCATCTTG 1080
DB 175544 AAGGCGCAGAAAGCGGAAATGAAATCCCTGATGAGTGGCGCTGTGCACTTTCATCTTG 175485
QY 1081 AAGCAGATTAATGACCGGATTAAGAGAGGCGCTGACAGTCTTGTTCACCGGGGAGAGGACAG 1140
DB 175484 AAGCAGATTAATGACCGGATTAAGAGAGGCGCTGACAGTCTTGTTCACCGGGGAGAGGACAG 175425
QY 1141 CTGACACTGGCGCTGAGCT 1157
DB 175424 CTGACACTGGCGCTGAGCT 175408

RESULT 10
AC009723 179592 bp DNA linear HTG 08-NOV-2000
LOCUS AC009723/c
DEFINITION Homo sapiens clone RP11-44F19, WORKING DRAFT SEQUENCE, 25 unordered
pieces.
AC009723.4 GI:9965551
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 179592)
Birken,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone RP11-44F19
Unpublished
2 (bases 1 to 179592)
Birken,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baker,J., Baldwin,J., Barna,N., Beckert,J., Benn,J., Brown,A.,
Castle,A., Cerny,J., Collingwood,M., Collins,S., Collymore,A.,
Cooke,P., DeRubeis,K., Depayre,E., Devore,K., Dewar,K.,
Donelan,L., Doyle,M., Ferreira,P., Fitzhugh,W., Forrest,C.,
Funke,R., Gage,D., Galagan,J., Gardy,J., Gilbert,D., Grant,G.,
Hagos,B., Heaford,A., Horton,L., Howland,J., Jones,C., Kahn,L.,
Karatas,A., Lehoczy,J., Lieu,C., Locke,K., MacDonald,P.,
Marguis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J.,
Melgrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
Naylor,J., Nifluff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,
Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,
Strange-Thomann,N., Stojanovic,N., Stone,C., Sudmanian,A.,
Tessier,S., Torrella-Miller,T., Vassiliev,H., Vo,A., Wagner,A.,
Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.
Direct Submission
Submitted (29-AUG-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 3, 2000 this sequence version replaced gi:7770419.
All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
```


Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L1062
Center clone name: 44.F.19
----- Summary Statistics
Sequencing vector: M13; M7815; 95% of reads
Sequencing vector: Plasmid; n/a; 5% of reads
Chemistry: Dye-terminator; Big Dye; 95% of reads
Chemistry: Dye-terminator; Big Dye; 95% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 159631 bases at least Q40
Consensus quality: 174012 bases at least Q30
Consensus quality: 174012 bases at least Q20
Insert size: 176000; agarose-IP
Insert size: 177192; sum-of-coverage
Quality coverage: 4.2 in Q20 bases.
NOTE: This is a 'working draft' sequence. It currently
* consists of 25 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 333: contig of 333 bp in length
3334 3433: gap of 100 bp
3434 4557: contig of 1124 bp in length
4558 4657: gap of 100 bp
4658 5689: contig of 1032 bp in length
5690 5789: gap of 100 bp
5790 7161: contig of 1372 bp in length
7162 7261: gap of 100 bp
7262 8676: contig of 1415 bp in length
8677 8776: gap of 100 bp
8777 9951: contig of 1175 bp in length
9952 10051: gap of 100 bp
10052 11777: contig of 1726 bp in length
11778 11877: gap of 100 bp
11878 13012: contig of 1135 bp in length
13013 13112: gap of 100 bp
13113 14573: contig of 1461 bp in length
14574 14673: gap of 100 bp
14674 17078: contig of 2405 bp in length
17079 17179: gap of 100 bp
17179 18544: contig of 1366 bp in length
18545 18644: gap of 100 bp
18645 20233: contig of 1589 bp in length
20234 20333: gap of 100 bp
20334 22479: contig of 2146 bp in length
22480 22579: gap of 100 bp
22580 24247: contig of 1668 bp in length
24248 24347: gap of 100 bp
24348 26991: contig of 2644 bp in length
26992 27091: gap of 100 bp
27092 31443: contig of 4352 bp in length
31444 31543: gap of 100 bp
31544 37218: contig of 5675 bp in length
37219 37318: gap of 100 bp
37319 40338: contig of 3020 bp in length
40339 40438: gap of 100 bp
40439 45330: contig of 4892 bp in length
45331 45430: gap of 100 bp
45431 53182: contig of 7752 bp in length
53183 53282: gap of 100 bp
53283 61147: contig of 7865 bp in length
61148 61247: gap of 100 bp
61249 73146: contig of 11699 bp in length
73147 73247: gap of 100 bp
73248 82900: contig of 9654 bp in length
82901 83000: gap of 100 bp

FEATURES
source
* 83001 139568: contig of 56568 bp in length
* 139569 139668: gap of 100 bp
* 139669 179592: contig of 39924 bp in length.
Location/Qualifiers
1. 179592
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="RP11-44F19"
/clone_id="RP11-11 Human Male BAC"
1. 3333
/note="assembly_fragment"
clone_end:56
vector_side:left
3434. 4557
/note="assembly_fragment"
4558. 5689
/note="assembly_fragment"
5790. 7161
/note="assembly_fragment"
7262. 8676
/note="assembly_fragment"
8777. 9951
/note="assembly_fragment"
10052. 11777
/note="assembly_fragment"
11878. 13012
/note="assembly_fragment"
13113. 14573
/note="assembly_fragment"
14674. 17078
/note="assembly_fragment"
17179. 18544
/note="assembly_fragment"
18645. 20233
/note="assembly_fragment"
20334. 22479
/note="assembly_fragment"
22580. 24247
/note="assembly_fragment"
24348. 26991
/note="assembly_fragment"
27092. 31443
/note="assembly_fragment"
31544. 37218
/note="assembly_fragment"
37319. 40338
/note="assembly_fragment"
40439. 45330
/note="assembly_fragment"
45431. 53182
/note="assembly_fragment"
53283. 61147
/note="assembly_fragment"
61248. 73146
/note="assembly_fragment"
73247. 82900
/note="assembly_fragment"
83001. 139568
/note="assembly_fragment"
139669. 179592
/note="assembly_fragment"
clone_end:77
vector_side:right

ORIGIN
Query Match 18.8%; Score 1068; DB 2; Length 179592;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1188; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 ATGAAAGCCATGCGCCGGAAGTGAAGTCTCTCTCCACCTCATGAGGCGCAG 60
Db 97654 ATGAAAGCCATGCGCCGGAAGTGAAGTCTCTCTCCACCTCATGAGGCGCAG 97595

```

QY 61 GGTCTCTCCACTTTGGTCAACCCGCGAGCCAGCCCGCTGTCCAGAGAGAGCGGTCAATT 120
Db 97594 GGTCTCTCCACTTTGGTCAACCCGCGAGCCAGCCCGCTGTCCAGAGAGAGCGGTCAATT 97535
QY 121 GTCAATTTCCGAGAGAGAGCCCGCGAGGGTTTCAATCAGCTGTGTGTATGAGAGACA 180
Db 97534 GTCAATTTCCGAGAGAGAGCCCGCGAGGGTTTCAATCAGCTGTGTATGAGAGACA 97475
QY 181 GGAACATTTACTTGGGGGCGGTCAATCGAATTTTACAGCTCTCCAGCAGCTGAAGGTC 240
Db 97474 GGAACATTTACTTGGGGGCGGTCAATCGAATTTTACAGCTCTCCAGCAGCTGAAGGTC 97415
QY 241 TTGGTACGATGAGAGAGAGCCCGCGAGAGAGACAACCCCAAGTTTACCAACCCCGCATC 300
Db 97414 TTGGTACGATGAGAGAGAGCCCGCGAGAGAGACAACCCCAAGTTTACCAACCCCGCATC 97355
QY 301 GTCCAGACTGCAATGAGCCCTGACCAACCAAGTCAAGAGATGCTCTCATATA 360
Db 97354 GTCCAGACTGCAATGAGCCCTGACCAACCAAGTCAAGAGATGCTCTCATATA 97295
QY 361 GACTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db 97294 GACTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 97235
QY 421 CTGAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Db 97234 CTGAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 97175
QY 481 TCAAGTGTCAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Db 97174 TCAAGTGTCAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 97115
QY 541 GACAAAGTGTTCATTTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db 97114 GACAAAGTGTTCATTTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 97055
QY 601 AGCCGAGAACTGACCAAGAACTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db 97054 AGCCGAGAACTGACCAAGAACTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 96995
QY 661 GAGTTCGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 719
Db 96994 GAGTTCGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 96935
QY 719 TGATATCTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 779
Db 96934 TGATATCTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 96875
QY 779 ACTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 839
Db 96874 ACTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 96815
QY 839 CGTGAAGTTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 899
Db 96814 CGTGAAGTTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 96755
QY 899 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 959
Db 96754 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 96695
QY 959 CGTGTCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1019
Db 96694 CGTGTCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 96635
QY 1019 CAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1079
Db 96634 CAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 96575
QY 1079 GAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1139
Db 96574 GAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 96515

```

```

QY 1140 GCTGAGACTGCGCTGCTGCTCAAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1188
Db 96514 GCTGAGACTGCGCTGCTGCTCAAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 96466

RESULT 11
HSM805617
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases, 1 to 3556)
Ostenwaeider, B., Obermaier, B., Newes, H. W., Gaassenhuber, J. and
Wiemann, S.
Direct Submission
Submitted (09-JUL-2002) 1, D-85764 Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
sequenced by Medigenomix (Martinsried/Germany) within the CDNA
sequencing consortium of the German Genome Project. This clone
(DKFZp56600546) is available at the RZPD in Berlin. Please contact
the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cdna/.
Location/Qualifiers
1..3556
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp56600546"
/db_xref="taxon:9606"
/clone="DKFZp56600546"
/tissue_type="kidney"
/clone_lib="566 (synonym: hfk2). Vector pAMP1; host
X1-blue; sites NotI + SalI"
/dev_stage="fetal"
1..3556
/gene="DKFZp56600546"
<1..1031
/gene="DKFZp56600546"
/notes="strong similarity to mouse plexin 2"
/codon_start=3
/product="hypothetical protein"
/protein_id="CAD39161.1"
/db_xref="GI:21740305"
/db_xref="SPTREMBL:Q8ND00"
/translation="LEWROGSGARNDIODEDITIKIENDMKRLNTIAHYVDPDSVVA
LVSKQVTAIVAVNNSTVSRISASKENMIRRTGSPDSIRSTPMITPLESGVMMHL
VKNHEHQKQKGRDSKRVSVSEIYLRRLATIGTLOKFPDDIPEFISIAHSGSLPLA
IKVMDFLDQADKRGHDPVIRHWNKSNCLPLRFWMDLNKQPFVFIHKSITDAG
LSVAQTFMDSCTSEHRLGKDSNKLIVKDVPSYKWNVERYSIDIGKPAISDQD
NNAVLAQOSRMHNEFNTMSALSEIFSVYKXSEIILQPLDHDQCGKQKLAYKLEQV
ITIMSLDS"
3532
/gene="DKFZp56600546"

polya_site

ORIGIN
Query Match 15.4%; Score 878; DB 9; Length 3556;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1028; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4655 ATCTGAGAGTGGGAGACAGAAATGTTGGGCAAGATGATCTTGAAGATGAAGCATCACCA 4714
Db 1 ATCTGAGAGTGGGAGACAGAAATGTTGGGCAAGATGATCTTGAAGATGAAGCATCACCA 60

QY 4715 CCAAGATGAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4774
Db 61 CCAAGATGAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120

```

4775 GTTCCGTTGGGATAGTGTGCAAGGATGACAGGCTATTAACGAGTGAACACTCCA 4834
Db 121 GTTCCGTTGGGATAGTGTGCAAGGATGACAGGCTATTAACGAGTGAACACTCCA 180
Qy 4835 CCGTCTCCAGAGCTCCAGCAAGTAATATGAAAAATGATCCGGTACACGGGACCCCG 4894
Db 181 CCGTCTCCAGAGCTCCAGCAAGTAATATGAAAAATGATCCGGTACACGGGACCCCG 240
Qy 4895 AACAGCTCCGCTCCAGAGCAACCTATGATCACTCTGTCCTGGAGAGTGGATCAAGATG 4954
Db 241 AACAGCTCCGCTCCAGAGCAACCTATGATCACTCTGTCCTGGAGAGTGGATCAAGATG 300
Qy 4955 GGCACTAGTGAAGAACCAAGAGCAGGAGACAGAAAGAGGGGGACCGGGGGAGCAAGA 5014
Db 301 GGCACTAGTGAAGAACCAAGAGCAGGAGACAGAAAGAGGGGGACCGGGGGAGCAAGA 360
Qy 5015 TGGTGTGGAATCTACTGACCCGACTCCCTGGGCACTAAGGGACACTGACAGATTTG 5074
Db 361 TGGTGTGGAATCTACTGACCCGACTCCCTGGGCACTAAGGGACACTGACAGATTTG 420
Qy 5075 TGGATGACCTCTTTGAGACCATCTTCAAGACGGGACACGGTGGCTTGCCTGCCCTGG 5134
Db 421 TGGATGACCTCTTTGAGACCATCTTCAAGACGGGACACGGTGGCTTGCCTGCCCTGG 480
Qy 5135 CCATCAAGTACATGTTTGAATCTTCTGGATGAGCAGGCTGATTAACATGGCATTCATGACC 5194
Db 481 CCATCAAGTACATGTTTGAATCTTCTGGATGAGCAGGCTGATTAACATGGCATTCATGACC 540
Qy 5195 CGCAGTCCGCAATACCTGGAAGAGCAATGCTGCCCTGAGGGTTTGGGTCACAATGA 5254
Db 541 CGCAGTCCGCAATACCTGGAAGAGCAATGCTGCCCTGAGGGTTTGGGTCACAATGA 600
Qy 5255 TCAGAAACCCGAGTGTGTGTTGATCATCCATAGAACAGATATACAGAGCCTGCTCT 5314
Db 601 TCAGAAACCCGAGTGTGTGTTGATCATCCATAGAACAGATATACAGAGCCTGCTCT 660
Qy 5315 CTGATGAGGCTCAGACCTTCATGAGACTTGTCTCCAGAGCAGGCTGGGCAAG 5374
Db 661 CTGATGAGGCTCAGACCTTCATGAGACTTGTCTCCAGAGCAGGCTGGGCAAG 720
Qy 5375 ACTGCGCTCCAAACAAGTGTGTATGCAAGAGCAATCCCAAGCTTACAAGATTTGGGTG 5434
Db 721 ACTGCGCTCCAAACAAGTGTGTATGCAAGAGCAATCCCAAGCTTACAAGATTTGGGTG 780
Qy 5435 AGAGTATTACTCAGACATAGAGAAATGACAGCATACAGCAACCAAGATGAGCAT 5494
Db 781 AGAGTATTACTCAGACATAGAGAAATGACAGCATACAGCAACCAAGATGAGCAT 840
Qy 5495 ACCTGGCTGAGCAATCCGAGTGCATGATGATGATGATGATGATGATGATGATGATG 5554
Db 841 ACCTGGCTGAGCAATCCGAGTGCATGATGATGATGATGATGATGATGATGATGATG 900
Qy 5555 AGATTTCTCTCTAGTGGGCAAAATACAGGAGAGAGATCTTGGACCTTGGACCAAGAG 5614
Db 901 AGATTTCTCTCTAGTGGGCAAAATACAGGAGAGAGATCTTGGACCTTGGACCAAGAG 960
Qy 5615 ACCAGTGTGGAAGAGCAAACTGGCTTCAAACTAGAAACAAGTATTAACCTCATAGAGT 5674
Db 961 ACCAGTGTGGAAGAGCAAACTGGCTTCAAACTAGAAACAAGTATTAACCTCATAGAGT 1020
Qy 5675 TAGACAGCTGA 5685
Db 1021 TAGACAGCTGA 1031

RESULT 12
AX574564 2597 bp DNA 1linear PAT 07-JAN-2003
LOCUS AX574564
DEFINITION Sequence 91 from Patent WO0224888.
ACCESSION AX574564
VERSION AX574564.1 GI:27551835
KEYWORDS

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 Baker, K.P., Eaton, D.L., Filvaroff, E., Goddard, A., Grimaldi, J.C., Gurney, A.L., Smith, V., Stephan, D.P., Matanabe, C.K., Wood, W.I., Zhang, Z. and Fong, S.
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding the same
JOURNAL Patent: WO 0224888-A 91 28-MAR-2002;
FEATURES
source Location/Qualifiers
1..2597
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 13.9%; Score 793; DB 6; Length 2597;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 579 CGAGTATTTCCACATCTCCAGCCGGAATCAACAAAGACTCTGAGCGGATGGCAT 638
Db 1 CGAGTATTTCCACATCTCCAGCCGGAATCAACAAAGACTCTGAGCGGATGGCAT 60
Qy 639 GTTGGCGTACGCTCTTCCATGATGATGATGATGATGATGATGATGATGATGATGATG 698
Db 61 GTTGGCGTACGCTCTTCCATGATGATGATGATGATGATGATGATGATGATGATGATG 120
Qy 699 CTTCAACATCAATCCCTGACCTTGAATATCTAATGATGATGATGATGATGATGATGATG 758
Db 121 CTTCAACATCAATCCCTGACCTTGAATATCTAATGATGATGATGATGATGATGATGATG 180
Qy 759 TGTCTACTTTTGAACCTTCCAACTGAGATGATGATGATGATGATGATGATGATGATGATG 818
Db 181 TGTCTACTTTTGAACCTTCCAACTGAGATGATGATGATGATGATGATGATGATGATGATG 240
Qy 819 GAGAGTATATACATCAAGCTCTGAGGCTTGGACAGAGGACAGAGCTTCACTCTCTA 878
Db 241 GAGAGTATATACATCAAGCTCTGAGGCTTGGACAGAGGACAGAGCTTCACTCTCTA 300
Qy 879 TGTAGAGTGTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 938
Db 301 TGTAGAGTGTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
Qy 939 CTAACGTGCAAGAGGGGGGGCGGTGTTGGAGAGACCTTGGAGTGCATCAAGATGATGA 998
Db 361 CTAACGTGCAAGAGGGGGGGCGGTGTTGGAGAGACCTTGGAGTGCATCAAGATGATGA 420
Qy 999 CCTGCTCTTCAACCGCTCTTCTCCAAAGGGCCAGAGCGAAATGAATCTCTGATGATG 1058
Db 421 CCTGCTCTTCAACCGCTCTTCTCCAAAGGGCCAGAGCGAAATGAATCTCTGATGATG 480
Qy 1059 GGCCCTGTGCAATCTTCAATCTTGAAGCAAGATTAAGACCGCATTAAGAGCGGCTGAC 1118
Db 481 GGCCCTGTGCAATCTTCAATCTTGAAGCAAGATTAAGACCGCATTAAGAGCGGCTGAC 540
Qy 1119 TTGTTACCGGGGCGAGGACAGCTGAGCTGAGCTGCTCAAGGTGAAGATCCCTG 1178
Db 541 TTGTTACCGGGGCGAGGACAGCTGAGCTGAGCTGCTCAAGGTGAAGATCCCTG 600
Qy 1179 CAGCAGTGGCTCTTAACATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 1238
Db 601 CAGCAGTGGCTCTTAACATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 660
Qy 1239 GGAAGTGTCCGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1298
Db 661 GGAAGTGTCCGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
Qy 1299 GTCTGTCACTGATATGCTTACAAAGACATCTCTTGGCTTTGTGGGACCAAAAGTGG 1358

Db 721 GTCTGTCATGTCATATGTCACAGAACCACTCTCTGAGCTTTGTGGGACCAAAAGTGG 780

Qy 1359 CAAGCTGAAGAG 1371

Db 781 CAAGCTGAAGAG 793

RESULT 13

LOCUS AY358850 2597 bp mRNA linear PRI 03-OCT-2003

DEFINITION Homo sapiens clone DNA190803 FAYV2820 (UNQ2820) mRNA, complete cds.

ACCESSION AY358850

VERSION AY358850.1 GI:37182817

KEYWORDS FLI CDNA.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 2597)

AUTHORS Clark,H.F., Gurney,A.L., Abaya,E., Baker,K., Baldwin,D., Brush,J., Chen,J., Chow,B., Chui,C., Crowley,C., Currell,B., Deuel,B., Dowd,P., Eaton,D., Foster,J., Grimaldi,C., Gu,Q., Hass,P.E., Heidens,S., Huang,A., Kim,H.S., Klimowski,L., Jin,Y., Johnson,S., Lee,J., Lewis,L., Liao,D., Mark,M., Robbie,E., Sanchez,C., Schoenfeld,J., Sehnagiri,S., Simmons,L., Singh,J., Smith,V., Stinson,J., Vagtes,A., Vandlen,R., Watanabe,C., Weand,D., Woods,K., Xie,M.H., Yansura,D., Yi,S., Yu,G., Yuan,J., Zhang,M., Zhang,Z., Goddard,A., Wood,W.I. and Godowski,P.

The Secreted Protein Discovery Initiative (SPDI), a large-scale effort to identify novel human secreted and transmembrane proteins: A Bioinformatics Assessment

JOURNAL Genome Res. 13 (10), 2265-2270 (2003)

PUBMED 12975309

REFERENCE 2 (bases 1 to 2597)

AUTHORS Clark,H.F.

TITLE Direct Submission

JOURNAL Submitted (01-AUG-2003) Department of Bioinformatics, Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080, USA

FEATURES

Source

1..2597

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="DNA190803"

1..2597

/locus_tag="UNQ2820"

59..901

/locus_tag="UNQ2820"

/note="PRO34003"

/codon_start=1

/product="FAYV2820"

/protein_id="AA089209.1"

/db_xref="GI:37182818"

/translation="MFATVFDHRYASMKIKPSTFTTTPIDYIYVGGSSGFYF LTLQEMWSPGSTKEQVYTSKVRLECKETANSVEVPIGERSGVERLLOAY LSKAGAVLGRIGVDDLLFTVFSKQKQKMSLDLSALCIPILKQINDRKERQ SCYREGSLDLMLKVDI PCSALTLTIDNFCGLDMNAPAGVSDVWRGIPVFERD RMTSVIAYVYKMSLAFVGTSGKSKKVPVSLCPTLELGTGPRSHRATVTEILFSS CSSN"

ORIGIN

Query Match 13.9%; Score 793; DB 9; Length 2597;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 579 CGAGATATTTCCACCATCTCCAGCGGAAATGACCAAGAACTCTGAGCGAGTGCAT 638

Db 1 CGAGATATTTCCACCATCTCCAGCGGAAATGACCAAGAACTCTGAGCGAGTGCAT 60

Qy 639 GTTCGGATAGCTCTTCATGATGAGTGTGCTGCTCGATGATTAAGATCCCTTGACAC 698

Db 61 GTTCGGATAGCTCTTCATGATGAGTGTGCTGCTCGATGATTAAGATCCCTTGACAC 120

Qy 699 CTTCACATCATCCCTGACTTGATATCTATGTTTGTAGAGTGGCACTT 758

Db 121 CTTCACATCATCCCTGACTTGATATCTATGTTTGTAGAGTGGCACTT 180

Qy 759 TGTCTACTTTTGGACCTTCACACCTGAGATGTGTCTCCACAGGCTCCACCAAGGA 818

Db 181 TGTCTACTTTTGGACCTTCACACCTGAGATGTGTCTCCACAGGCTCCACCAAGGA 240

Qy 819 GCAGGTATATACATCAAGCTCGTAGAGCTTTGCAAGAGACACAGCCTTCACTCTTA 878

Db 241 GCAGGTATATATCATCAAGCTCGTAGAGCTTTGCAAGAGACACAGCCTTCACTCTTA 300

Qy 879 TGTAGAGTGGCCCATTTGGCTGTGACCGCAGATGGGGGTGAGTACCCCTGCTGCAGGCTGC 938

Db 301 TGTAGAGTGGCCCATTTGGCTGTGACCGCAGATGGGGGTGAGTACCCCTGCTGCAGGCTGC 360

Qy 939 CTACCTGTCCAAAGCGGGGCGGTGCTTGACAGACCTTGAGTCCATCCAGATGATGA 998

Db 361 CTACCTGTCCAAAGCGGGGCGGTGCTTGACAGACCTTGAGTCCATCCAGATGATGA 420

Qy 999 CCTGCTCTTACCGTCTTCTCCAGAGGCGCAGAGCGGAAATGAAATCCCTGATGAGTGC 1058

Db 421 CCTGCTCTTACCGTCTTCTCCAGAGGCGCAGAGCGGAAATGAAATCCCTGATGAGTGC 480

Qy 1059 GGCCCTGTGCATCTTCATCTGAAAGATTAATGACCGCATTAAGAGCGGCTGACATC 1118

Db 481 GGCCCTGTGCATCTTCATCTGAAAGATTAATGACCGCATTAAGAGCGGCTGACATC 540

Qy 1119 TTGTTACCGGGGCGAGGGCAGCGCTGACCTGCGCTGCTCAAGTGAAGACATCCCTG 1178

Db 541 TTGTTACCGGGGCGAGGGCAGCGCTGACCTGCGCTGCTCAAGTGAAGACATCCCTG 600

Qy 1179 CAGCAATGGGCTCTTAACATTAAGATTAATCTTGTGTGGCCCTGACATGAATGCTCCCT 1238

Db 601 CAGCAATGGGCTCTTAACATTAAGATTAATCTTGTGTGGCCCTGACATGAATGCTCCCT 660

Qy 1239 GGGAGTGTCCGACATGATGTCGTGATTTCCCTCTTCAAGAGACAGGAGCCGATGAC 1298

Db 661 GGGAGTGTCCGACATGATGTCGTGATTTCCCTCTTCAAGAGACAGGAGCCGATGAC 720

Qy 1299 GTCTGTCATGTCATATGTCATCAAGAACATCTCTGCGCTTTGGGACCAAAAGTGG 1358

Db 721 GTCTGTCATGTCATATGTCATCAAGAACATCTCTGCGCTTTGGGACCAAAAGTGG 780

Qy 1359 CAAGCTGAAGAG 1371

Db 781 CAAGCTGAAGAG 793

RESULT 14

LOCUS HSM802037 1029 bp mRNA linear PRI 18-FEB-2000

DEFINITION Homo sapiens mRNA, cDNA DKFZp434G0625 (from clone DKFZp434G0625); partial cds.

ACCESSION AL137352

VERSION AL137352.1 GI:6807867

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1029)

AUTHORS Amstutz,W., Winkler,U., Mewes,H.W., Gassenhuber,J. and Wiemann,S.

TITLE Direct Submission

JOURNAL Submitted (15-JAN-2000) MIPS, Am Klopferplatz 18a, D-82152 Martinsried, GERMANY

COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by EMBL (European Molecular Biology Laboratories, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.

This clone (DKFZp434G0625) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059


```
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
1
727: contig of 727 bp in length
728
827: gap of 100 bp
828
1590: contig of 763 bp in length
1591
1690: gap of 100 bp
1691
2443: contig of 753 bp in length
2444
2543: gap of 100 bp
2544
3286: contig of 743 bp in length
3287
3386: gap of 100 bp
3387
4099: contig of 713 bp in length
4100
4199: gap of 100 bp
4200
4924: contig of 725 bp in length
4925
5024: gap of 100 bp
5025
5745: contig of 721 bp in length
5746
5845: gap of 100 bp
5846
6667: contig of 722 bp in length
6668
6667: gap of 100 bp
6668
7409: contig of 742 bp in length
7410
7509: gap of 100 bp
7510
8339: contig of 730 bp in length
8340
9048: gap of 100 bp
9049
9149
9902: contig of 754 bp in length
9903
10002: gap of 100 bp
10003
10745: contig of 743 bp in length
10746
10845: gap of 100 bp
10846
11593: contig of 748 bp in length
11594
11693: gap of 100 bp
11694
12426: contig of 733 bp in length
12427
12526: gap of 100 bp
12527
13240: contig of 714 bp in length
13241
13341
14074: gap of 100 bp
14074: contig of 734 bp in length
14174: gap of 100 bp
14175
14806: contig of 732 bp in length
14807
15006: gap of 100 bp
15007
15728: contig of 722 bp in length
15729
15828: gap of 100 bp
15829
16542: contig of 714 bp in length
16543
16642: gap of 100 bp
16643
17379: contig of 737 bp in length
17380
17479: gap of 100 bp
17480
18225: contig of 746 bp in length
18226
18325: gap of 100 bp
18326
19056: contig of 731 bp in length
19057
19156: gap of 100 bp
19157
19894: contig of 738 bp in length
19895
19994: gap of 100 bp
19995
20729: contig of 735 bp in length
20730
20829: gap of 100 bp
20830
21569: contig of 740 bp in length
21570
21669: gap of 100 bp
21670
22401: contig of 732 bp in length
22402
22501: gap of 100 bp
22502
23325: contig of 724 bp in length
23326
23325: gap of 100 bp
23326
24053: contig of 728 bp in length
24054
24153: gap of 100 bp
24154
24892: contig of 733 bp in length
24893
24992: gap of 100 bp
24993
25727: contig of 735 bp in length
25728
25827: gap of 100 bp
25828
26566: contig of 733 bp in length
26567
26666: gap of 100 bp
26667
27410: contig of 744 bp in length
27411
27510: gap of 100 bp
27511
28341: contig of 731 bp in length
28342
28342: gap of 100 bp
28343
29090: contig of 749 bp in length
29091
29190: gap of 100 bp
29191
29923: contig of 733 bp in length
29924
30023: gap of 100 bp
30024
30742: contig of 719 bp in length
30743
30842: gap of 100 bp
30843
31558: contig of 716 bp in length
31559
31658: gap of 100 bp
31659
32388: contig of 730 bp in length
32389
32488: gap of 100 bp
32489
33204: contig of 716 bp in length
33205
33304: gap of 100 bp
33305
34018: contig of 714 bp in length
34019
34118: gap of 100 bp
34119
34850: contig of 732 bp in length
34851
34950: gap of 100 bp
34951
35689: contig of 739 bp in length
35690
35789: gap of 100 bp
35790
36536: contig of 747 bp in length
36537
36636: gap of 100 bp
36637
37354: contig of 718 bp in length
37355
37454: gap of 100 bp
37455
38198: contig of 744 bp in length
38199
38298: gap of 100 bp
38299
39045: contig of 747 bp in length
39046
39145: gap of 100 bp
39146
39872: contig of 727 bp in length
39873
39972: gap of 100 bp
39973
40679: contig of 707 bp in length
40680
40779: gap of 100 bp
40780
41511: contig of 732 bp in length
41512
41611: gap of 100 bp
41612
42342: contig of 731 bp in length
42343
42442: gap of 100 bp
42443
43178: contig of 736 bp in length
43179
43278: gap of 100 bp
43279
44022: contig of 744 bp in length
44023
44122: gap of 100 bp
44123
44860: contig of 738 bp in length
44861
44960: gap of 100 bp
44961
45676: contig of 716 bp in length
45677
45776: gap of 100 bp
45777
46315: contig of 735 bp in length
46316
46415: gap of 100 bp
46416
47350: contig of 735 bp in length
47351
47450: gap of 100 bp
47451
48148: contig of 698 bp in length
48149
48248: gap of 100 bp
48249
48949: contig of 701 bp in length
48950
49049: gap of 100 bp
49050
49775: contig of 726 bp in length
49776
49875: gap of 100 bp
49876
50592: contig of 717 bp in length
50593
50692: gap of 100 bp
50693
51419: contig of 727 bp in length
51420
51519: gap of 100 bp
51520
52272: contig of 753 bp in length
52273
52372: gap of 100 bp
52373
53123: contig of 751 bp in length
53124
53223: gap of 100 bp
53224
53965: contig of 742 bp in length
53966
54065: gap of 100 bp
54066
54780: contig of 715 bp in length
54781
54880: gap of 100 bp
54881
55621: contig of 741 bp in length
55622
55721: gap of 100 bp
55722
56458: contig of 737 bp in length
56459
56558: gap of 100 bp
```

* 56559 57291: contig of 733 bp in length
* 57292 57391: gap of 100 bp

Query Match 8.9%; Score 509; DB 2; Length 74655;
Best Local Similarity 99.8%; Pred. No. 9.9e-283;
Matches 559; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 94 CCGCTGCCAGAGAGCGGTCATTGTCACATCCGAGAGAGAGCCGCCGAGGGTTTC 153
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 33970 CCGCTGCCAGAGAGCGGTCATTGTCACATCCGAGAGAGAGCCGCCGAGGGTTTC 33911
QY 154 AATCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 213
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 33910 AATCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 33851
QY 214 TACAGCTCTCCAGCGGCTGTGAAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGT 273
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 33850 TACAGCTCTCCAGCGGCTGTGAAGCTTGTGTGTGTGTGTGTGTGTGTGTGT 33791
QY 274 AACCCCAAGTGTACCCAGCCCGCATGTCAGACTGCAATGAGCCCTGACCAACC 333
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 33790 AACCCCAAGTGTACCCAGCCCGCATGTCAGACTGCAATGAGCCCTGACCAACC 33731
QY 334 AACATGTCAACAGATGCTCTCATGACTACAGGAGAACAGGCTGATTGCTGTGG 393
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 33730 AACATGTCAACAGATGCTCTCATGACTACAGGAGAACAGGCTGATTGCTGTGG 33671
QY 394 AGCTGTACCAAGGATCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAG 453
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 33670 AGCTGTACCAAGGATCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAG 33611
QY 454 CTTTATATTAAGAGAGCACTATCTGTCAAGGTGTCAAGAGAGCGGCTGAGTTTGA 513
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 33610 CTTTATATTAAGAGAGCACTATCTGTCAAGGTGTCAAGAGAGCGGCTGAGTTTGA 33551
QY 514 GTGATGCTCTCTCAGAGACCTGATGACAAAGCTGTTCAATTGCAAGCGAGTGGG 573
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 33550 GTGATGCTCTCTCAGAGACCTGATGACAAAGCTGTTCAATTGCAAGCGAGTGGG 33491
QY 574 AAGCCCGAGTATTTTCCACCATCTCCAGCCGGAACAGCAAGAACTTGAAGCGGAT 633
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 33490 AAGCCCGAGTATTTTCCACCATCTCCAGCCGGAACAGCAAGAACTTGAAGCGGAT 33431
QY 634 GGCAATGTCGCGTACGCTT 653
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 33430 GGCAATGTCGCGTACGCTT 33411
```

Search completed: February 20, 2004, 07:08:38
Job time : 14204 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 23:14:05 ; Search time 1340 Seconds
(without alignments)
18042.161 Million cell updates/sec

Title: US-09-964-956-12
Perfect score: 5691
Sequence: 1 atgaagccatgcctcgaa.....gcttagacagctgaataaa 5691

Scoring table: OLIGO-NTC
Gapop 60.0 , Gapext 60.0

Searched: 3373863 seqs, 212409041 residues

Word size : 0
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : N_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2002as:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2003cs:*
9: Geneseq2004s:*
10: Geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5691	100.0	5691	6	ABSG6437 Human CDN
2	5332	97.2	6367	6	ABR00062 CADHP-9 c
3	793	13.9	2597	6	ABK70006 CDNA enco
4	793	13.9	2597	8	ADA01359 Human PRO
5	793	13.9	2597	8	ADA43788 Human CDN
6	793	13.9	2597	8	ADA43556 Human CDN
7	793	13.9	2597	8	ADA01231 Human PRO
8	793	13.9	2597	8	ADA01115 Human CDN
9	793	13.9	2597	8	ADA43672 Human CDN
10	793	13.9	2597	8	ADA06934 Human PRO
11	793	13.9	2597	8	ADA08422 Novel hum
12	793	13.9	2597	8	ADB99715 Human PRO
13	793	13.9	2597	8	ADB86998 Human PRO
14	793	13.9	2597	8	ADB66153 Human CDN
15	793	13.9	2597	9	ADB99831 Human PRO
16	793	13.9	2597	9	ADB99486 Novel hum
17	793	13.9	2597	9	ADB66037 Human CDN
18	793	13.9	2597	9	ADC23435 Human CDN
19	793	13.9	2597	9	ADC26128 Human PRO
20	793	13.9	2597	9	ADB04955 Human PRO
21	793	13.9	2597	9	ADB11261 Human PRO
22	793	13.9	2597	9	ADB88192 Human PRO
23	793	13.9	2597	9	ADB95487 Human CDN

24	793	13.9	2597	9	ADB06417 Human PRO
25	793	13.9	2597	9	ADB38192 Human PRO
26	793	13.9	2597	9	ADB88308 Human PRO
27	793	13.9	2597	9	ADB90889 Human CDN
28	793	13.9	2597	10	ADB51742 Human CDN
29	793	13.9	2597	10	ADB51858 Human CDN
30	793	13.9	2597	10	ADB37716 Human CDN
31	793	13.9	2597	10	ADB37600 Human CDN
32	793	13.9	2597	10	ADB95371 Human CDN
33	793	13.9	2597	10	ADB38071 Human PRO
34	793	13.9	2597	10	ADB76160 Human PRO
35	793	13.9	2597	10	ADB39483 Human PRO
36	793	13.9	2597	10	ADB04287 Human PRO
37	793	13.9	2597	10	ADB39884 Human PRO
38	793	13.9	2597	10	ADB19749 Human CDN
39	793	13.9	2597	10	ADB77327 Human CDN
40	793	13.9	2597	10	ADB65435 Human PRO
41	793	13.9	2597	10	ADB76044 Human PRO
42	793	13.9	2597	10	ADB37955 Human PRO
43	793	13.9	2597	10	ADB64565 Human PRO
44	793	13.9	2597	10	ADB38900 Human PRO
45	793	13.9	2597	10	ADB51974 Human CDN
46	793	13.9	2597	10	ADB91005 Human CDN
47	793	13.9	2597	10	ADB38784 Human PRO
48	793	13.9	2597	10	ADB37484 Human CDN
49	793	13.9	2597	10	ADB06301 Human PRO
50	793	13.9	2597	10	ADB90160 Human CDN
51	793	13.9	2597	10	ADB38668 Human PRO
52	793	13.9	2597	10	ADB39599 Human PRO
53	793	13.9	2597	10	ADB89204 Human PRO
54	793	13.9	2597	10	ADB88971 Human PRO
55	793	13.9	2597	10	ADB19865 Human PRO
56	793	13.9	2597	10	ADB77443 Human CDN
57	793	13.9	2597	10	ADB63319 Human PRO
58	793	13.9	2597	10	ADB89971 Human PRO
59	793	13.9	2597	10	ADB38552 Human CDN
60	503	8.8	598	4	AA142745 Probe #11
61	503	8.8	598	4	AAK36939 Human bon
62	503	8.8	598	4	AAK11122 Human bra
63	503	8.8	598	4	ABSG6411 Human liv
64	491	8.6	601	4	AA800023 Human CDN
65	481	8.5	495	4	AA155806 Probe #24
66	481	8.5	495	4	AAK49882 Human bon
67	481	8.5	495	4	AAK23792 Human bra
68	481	8.5	495	4	AB849520 Human liv
69	479	8.4	3560	9	ADB62416 Human CDN
70	185	3.3	477	5	AA868294 DNA enco
71	185	3.3	477	5	ACD05873 Novel hum
72	183	3.2	1314	5	AA868296 DNA enco
73	183	3.2	1314	7	ACD05552 DNA enco
74	183	3.2	1314	3	ACD05552 DNA enco
75	183	3.2	1314	3	ACD05552 DNA enco
76	183	3.2	1314	3	ACD05552 DNA enco
77	183	3.2	1314	3	ACD05552 DNA enco
78	183	3.2	1314	3	ACD05552 DNA enco
79	183	3.2	1314	3	ACD05552 DNA enco
80	183	3.2	1314	3	ACD05552 DNA enco
81	183	3.2	1314	3	ACD05552 DNA enco
82	183	3.2	1314	3	ACD05552 DNA enco
83	183	3.2	1314	3	ACD05552 DNA enco
84	183	3.2	1314	3	ACD05552 DNA enco
85	183	3.2	1314	3	ACD05552 DNA enco
86	183	3.2	1314	3	ACD05552 DNA enco
87	183	3.2	1314	3	ACD05552 DNA enco
88	183	3.2	1314	3	ACD05552 DNA enco
89	183	3.2	1314	3	ACD05552 DNA enco
90	183	3.2	1314	3	ACD05552 DNA enco
91	183	3.2	1314	3	ACD05552 DNA enco
92	183	3.2	1314	3	ACD05552 DNA enco
93	183	3.2	1314	3	ACD05552 DNA enco
94	183	3.2	1314	3	ACD05552 DNA enco
95	183	3.2	1314	3	ACD05552 DNA enco
96	183	3.2	1314	3	ACD05552 DNA enco

C 97	23	0.4	190	3	AAC16570	Human sec	C 170	20	0.4	38506	7	ADA09418	
C 98	23	0.4	201	4	AA166420	Probe #25	C 171	20	0.4	88421	6	AA140781	
C 99	23	0.4	201	6	ABS23899	Human gen	C 172	20	0.4	110000	7	ABX16390_5	
C 100	23	0.4	600	4	AA143440	Probe #12	C 173	19	0.3	132	3	AAA41395	
C 101	23	0.4	600	6	ABS11541	Human gen	C 174	19	0.3	294	4	ACD92078	
C 102	22	0.4	22	6	ABS63466	Real-time	C 175	19	0.3	360	6	AA517588	
C 103	21	0.4	21	6	ABS63466	Real-time	C 176	19	0.3	436	6	AA656958	
C 104	21	0.4	154	5	ABV04995	Human pro	C 177	19	0.3	464	8	ACH44098	
C 105	21	0.4	205	5	ABV35258	Human pro	C 178	19	0.3	503	5	AAF59651	
C 106	21	0.4	205	5	ABV44095	Human pro	C 179	19	0.3	512	5	AA585908	
C 107	21	0.4	522	5	ABV14164	Human pro	C 180	19	0.3	531	2	AA701576	
C 108	21	0.4	3111	4	AA504276	Rat gluta	C 181	19	0.3	558	5	AAH97972	
C 109	21	0.4	3111	9	ADBS8570	Toxicity-	C 182	19	0.3	588	2	AAV22147	
C 110	21	0.4	3111	9	ADBS3197	Primary r	C 183	19	0.3	689	3	AAF08602	
C 111	21	0.4	135638	7	ABX34289	S. atrocol	C 184	19	0.3	720	5	AA574665	
C 112	20	0.4	274	6	ABL72690	Corn taas	C 185	19	0.3	735	7	ACF67304	
C 113	20	0.4	311	7	ABX88561	Corn ear-	C 186	19	0.3	739	4	AAK92012	
C 114	20	0.4	318	3	AAZ42602	Human 5'	C 187	19	0.3	739	4	AAK93668	
C 115	20	0.4	324	6	ABN94554	Gene #105	C 188	19	0.3	791	6	AA138167	
C 116	20	0.4	328	4	AA577777	CDNA #453	C 189	19	0.3	900	2	AAQ61558	
C 117	20	0.4	345	5	AA582885	DNA enco	C 190	19	0.3	1059	6	ABQ42478	
C 118	20	0.4	369	8	ACH21424	Human adu	C 191	19	0.3	1059	6	ABQ42479	
C 119	20	0.4	382	3	AACT07272	Human sec	C 192	19	0.3	1059	6	ABQ20484	
C 120	20	0.4	386	8	ACH29631	Human tes	C 193	19	0.3	1059	6	ABQ20485	
C 121	20	0.4	412	3	AACT07271	Human sec	C 194	19	0.3	1176	5	AA576846	
C 122	20	0.4	414	8	ACH29208	Human adu	C 195	19	0.3	1233	7	ACA21173	
C 123	20	0.4	419	6	ABK45224	CDNA enco	C 196	19	0.3	1236	8	AD332039	
C 124	20	0.4	456	3	AACT07270	Human sec	C 197	19	0.3	1296	6	AB232409	
C 125	20	0.4	462	3	AAAT7885	CDNA enco	C 198	19	0.3	1317	6	AA562393	
C 126	20	0.4	462	4	AA128623	Colon tum	C 199	19	0.3	1467	7	AA558328	
C 127	20	0.4	462	7	AB232809	Human col	C 200	19	0.3	1599	9	AD507844	
C 128	20	0.4	476	8	ACH44821	Human foe	C 201	19	0.3	2000	7	AD472625	
C 129	20	0.4	482	8	ACH33668	Human end	C 202	19	0.3	2036	4	AB111823	
C 130	20	0.4	486	3	AACT07273	Human sec	C 203	19	0.3	2206	4	AA502663	
C 131	20	0.4	495	8	ACH50755	Human nam	C 204	19	0.3	2206	7	AB267186	
C 132	20	0.4	500	7	ACD93147	Human col	C 205	19	0.3	2206	7	AB273589	
C 133	20	0.4	506	7	ADD19817	Seabass p	C 206	19	0.3	2206	7	ADA98098	
C 134	20	0.4	615	3	AACT00732	Human sec	C 207	19	0.3	2237	2	AAQ89779	
C 135	20	0.4	616	3	AAAI6233	Human col	C 208	19	0.3	2414	4	AAK94356	
C 136	20	0.4	670	3	AACT10893	Human sec	C 209	19	0.3	2444	7	AD542653	
C 137	20	0.4	689	3	AACT10890	Human sec	C 210	19	0.3	2466	9	ADBS8952	
C 138	20	0.4	700	6	ABN96801	Gene #129	C 211	19	0.3	2466	9	ADBS51694	
C 139	20	0.4	736	5	AA569768	DNA enco	C 212	19	0.3	2477	2	AA532244	
C 140	20	0.4	744	3	AACT10882	Human sec	C 213	19	0.3	2477	3	AA585855	
C 141	20	0.4	746	5	AA569769	DNA enco	C 214	19	0.3	2477	3	AAAT77558	
C 142	20	0.4	750	3	AAAI01845	Human col	C 215	19	0.3	2477	3	ADCT6489	
C 143	20	0.4	792	6	ABQ56140	Human ova	C 216	19	0.3	2477	4	AA572402	
C 144	20	0.4	824	4	AAH05487	Human CDN	C 217	19	0.3	2477	4	AA521409	
C 145	20	0.4	1320	4	AA56182	Salmonell	C 218	19	0.3	2477	7	ACA519035	
C 146	20	0.4	1320	7	ACAI51334	Prokaryot	C 219	19	0.3	2477	7	ACD24018	
C 147	20	0.4	1636	2	AACT39794	Mouse SH3	C 220	19	0.3	2477	7	ACA58432	
C 148	20	0.4	2060	4	AAH15894	Human CDN	C 221	19	0.3	2477	7	ACA67159	
C 149	20	0.4	2074	3	AACT6166	Human ORF	C 222	19	0.3	2477	7	ACA60139	
C 150	20	0.4	2153	4	AA160597	Human pol	C 223	19	0.3	2477	7	ACD07539	
C 151	20	0.4	2214	6	ABK34791	Human CDN	C 224	19	0.3	2477	7	ACA03768	
C 152	20	0.4	2444	3	AAA90083	Human neg	C 225	19	0.3	2477	7	ABX71587	
C 153	20	0.4	2522	2	AA158811	Human pol	C 226	19	0.3	2477	7	ACH06919	
C 154	20	0.4	2522	8	ADBA48792	Novel hum	C 227	19	0.3	2477	7	ABX89306	
C 155	20	0.4	2564	9	ADBS3125	Human CDN	C 228	19	0.3	2477	7	ACD41960	
C 156	20	0.4	2971	4	AAK93432	Human CDN	C 229	19	0.3	2477	7	ABX96156	
C 157	20	0.4	3849	4	AA525795	S. chryso	C 230	19	0.3	2477	7	ACA05477	
C 158	20	0.4	6252	4	AA500020	Human CDN	C 231	19	0.3	2477	7	ACD20144	
C 159	20	0.4	6252	5	AAH81745	Human dif	C 232	19	0.3	2477	7	ACA04189	
C 160	20	0.4	6329	6	ABL67279	Thyroid c	C 233	19	0.3	2477	7	ACA54947	
C 161	20	0.4	6329	4	AAH98400	Human EST	C 234	19	0.3	2477	8	ACD19782	
C 162	20	0.4	6329	4	ABAC9065	Human p1e	C 235	19	0.3	2477	8	ADA45850	
C 163	20	0.4	6561	5	AA585204	DNA enco	C 236	19	0.3	2477	8	ADA76281	
C 164	20	0.4	13882	3	AAZ87297	S. venezu	C 237	19	0.3	2477	8	ADB29374	
C 165	20	0.4	36778	3	AAZ87318	S. venezu	C 238	19	0.3	2477	8	ADA18931	
C 166	20	0.4	37948	3	AAZ87285	S. venezu	C 239	19	0.3	2477	8	ADA61554	
C 167	20	0.4	38506	3	AAAT75633	Nucleotid	C 240	19	0.3	2477	8	ADB19339	
C 168	20	0.4	38506	3	AAZ56001	Recombina	C 241	19	0.3	2477	8	ADB27880	
C 169	20	0.4	38506	7	AB556090	S. venezu	C 242	19	0.3	2477	8	ADA86359	
AdA09418 Cosmid pk							AdA09418						
AdA10781 88421nt g							AdA10781						
Continuation (6 of							AdA16390_5						
AdA41395 Human sec							AdA41395						
AdC92078 Human col							AdC92078						
AdA17588 DNA enco							AdA17588						
AdC65658 Human mic							AdC65658						
AdA40698 Human foe							AdA40698						
AdA59651 Fei ghrel							AdA59651						
AdA65908 DNA enco							AdA65908						
AdA01576 Human leu							AdA01576						
AdA97972 Murine 7-							AdA97972						
AdA22147 Mouse BH3							AdA22147						
AdA08602 Fusarium							AdA08602						
AdA74665 DNA enco							AdA74665						
AdA67304 Photohab							AdA67304						
AdA92012 Human CDN							AdA92012						
AdA93668 Human CDN							AdA93668						
AdA138167 DNA from							AdA138167						
AdA61558 Genomic s							AdA61558						
AdA42478 Oligonuc1							AdA42478						
AdA20484 Oligonuc1							AdA20484						
AdA20485 Oligonuc1							AdA20485						
AdA76846 DNA enco							AdA76846						
AdA21173 Prokaryot							AdA21173						
AdA33409 DNA enco							AdA33409						
AdA232409 Candida a							AdA232409						
AdA56393 CDNA sequ							AdA56393						
AdA558293 CDNA sequ							AdA558293						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ							AdA558282						
AdA558282 CDNA sequ													

243	19	0.3	2477	8	ADB15923	Adb15923	Human	PRO	316	19	0.3	2477	9	ADB74831	Adb74831	Human	sec
244	19	0.3	2477	8	ADA47709	Ada47709	Human	sec	317	19	0.3	2477	9	ADB47079	Adb47079	Human	hum
245	19	0.3	2477	8	ADA18230	Ada18230	Human	sec	318	19	0.3	2477	9	ADB66686	Adb66686	Human	PRO
246	19	0.3	2477	8	ACD66929	Acde6929	Human	CDN	319	19	0.3	2477	9	ADB77291	Adb77291	Novel	hum
247	19	0.3	2477	8	ADA67504	Ada67504	Human	PRO	320	19	0.3	2477	9	ADB34448	Adb34448	Human	PRO
248	19	0.3	2477	8	ADBS0511	Adbs0511	CDNA	enco	321	19	0.3	2477	9	ADBS5552	Adbs5552	Human	PRO
249	19	0.3	2477	8	ADBS8807	Adbs8807	Novel	hum	322	19	0.3	2477	9	ADBS3896	Adbs3896	Human	PRO
250	19	0.3	2477	8	ADA97019	Ada97019	Human	PRO	323	19	0.3	2477	9	ADBS5000	Adbs5000	Human	PRO
251	19	0.3	2477	8	ADA79323	Ada79323	Human	PRO	324	19	0.3	2477	9	ADBS6104	Adbs6104	Human	PRO
252	19	0.3	2477	8	ADA87462	Ada87462	Novel	hum	325	19	0.3	2477	9	ADB46499	Adb46499	Novel	hum
253	19	0.3	2477	8	ADBI6664	Adbi6664	Human	PRO	326	19	0.3	2477	9	ADC39677	Adc39677	Human	sec
254	19	0.3	2477	8	ACD83090	Acdb83090	Human	PRO	327	19	0.3	2477	9	ADC40191	Adc40191	Human	sec
255	19	0.3	2477	8	ADA16205	Ada16205	Human	sec	328	19	0.3	2477	9	ADC19015	Adc19015	Human	sec
256	19	0.3	2477	8	ADA91756	Ada91756	Novel	hum	329	19	0.3	2477	9	ADC34315	Adc34315	Human	sec
257	19	0.3	2477	8	ADBI4819	Adbi4819	Human	PRO	330	19	0.3	2477	9	ADC29370	Adc29370	Human	sec
258	19	0.3	2477	8	ADBI8780	Adbi8780	Novel	hum	331	19	0.3	2477	9	ADC28901	Adc28901	Human	sec
259	19	0.3	2477	8	ADA93995	Ada93995	Human	PRO	332	19	0.3	2477	9	ADC40786	Adc40786	Human	sec
260	19	0.3	2477	8	ADBI19891	Adbi19891	Novel	hum	333	19	0.3	2477	9	ADC19443	Adc19443	Human	sec
261	19	0.3	2477	8	ADBI13203	Adbi13203	Human	PRO	334	19	0.3	2477	9	ADC33891	Adc33891	Human	sec
262	19	0.3	2477	8	ACD98589	Acdb98589	Novel	hum	335	19	0.3	2477	9	ADC12961	Adc12961	Human	sec
263	19	0.3	2477	8	ADA74457	Ada74457	Human	PRO	336	19	0.3	2477	9	ADC50372	Adc50372	Novel	hum
264	19	0.3	2477	8	ADA42350	Ada42350	Human	sec	337	19	0.3	2477	9	ADC71919	Adc71919	Novel	hum
265	19	0.3	2477	8	ADB24690	Adb24690	Human	PRO	338	19	0.3	2477	9	ADC59898	Adc59898	Novel	hum
266	19	0.3	2477	8	ADA82214	Ada82214	Human	PRO	339	19	0.3	2477	9	ADC52905	Adc52905	Novel	hum
267	19	0.3	2477	8	ADA75177	Ada75177	Human	PRO	340	19	0.3	2477	9	ADC57259	Adc57259	Novel	hum
268	19	0.3	2477	8	ADA85255	Ada85255	Novel	hum	341	19	0.3	2477	9	ADC60450	Adc60450	Novel	hum
269	19	0.3	2477	8	ADA84703	Ada84703	Novel	hum	342	19	0.3	2477	9	ADC50925	Adc50925	Novel	hum
270	19	0.3	2477	8	ACD23268	Acdb23268	Human	PRO	343	19	0.3	2477	9	ADC65452	Adc65452	Human	PRO
271	19	0.3	2477	8	ADB29959	Adb29959	CDNA	enco	344	19	0.3	2477	9	ADC54550	Adc54550	Novel	hum
272	19	0.3	2477	8	ADA80487	Ada80487	Human	PRO	345	19	0.3	2477	9	ADC54550	Adc54550	Novel	hum
273	19	0.3	2477	8	ADA75729	Ada75729	Human	PRO	346	19	0.3	2477	9	ADC53511	Adc53511	Novel	hum
274	19	0.3	2477	8	ADA64954	Ada64954	Human	PRO	347	19	0.3	2477	9	ADC59034	Adc59034	Novel	hum
275	19	0.3	2477	8	ADB25250	Adb25250	Human	PRO	348	19	0.3	2477	9	ADC58482	Adc58482	Novel	hum
276	19	0.3	2477	8	ADA93426	Ada93426	Human	PRO	349	19	0.3	2477	9	ADC12413	Adc12413	Human	sec
277	19	0.3	2477	8	ADB26776	Adb26776	CDNA	enco	350	19	0.3	2477	9	ADC03156	Adc03156	Novel	hum
278	19	0.3	2477	8	ADBI1063	Adbi1063	CDNA	enco	351	19	0.3	2477	9	ADC90148	Adc90148	Novel	hum
279	19	0.3	2477	8	ADA60991	Ada60991	Homo sapi		352	19	0.3	2477	9	ADC69567	Adc69567	CDNA	enco
280	19	0.3	2477	8	ADB24138	Adb24138	Human	PRO	353	19	0.3	2477	9	ADC48456	Adc48456	Human	PRO
281	19	0.3	2477	8	ADA96467	Ada96467	Human	PRO	354	19	0.3	2477	9	ADC09985	Adc09985	Human	PRO
282	19	0.3	2477	8	ADA81039	Ada81039	Human	PRO	355	19	0.3	2477	9	ADC04560	Adc04560	Novel	hum
283	19	0.3	2477	8	ADA95915	Ada95915	Human	PRO	356	19	0.3	2477	9	ADC80516	Adc80516	Novel	hum
284	19	0.3	2477	8	ADB26224	Adb26224	CDNA	enco	357	19	0.3	2477	9	ADC11023	Adc11023	Human	PRO
285	19	0.3	2477	8	ADB21709	Adb21709	Novel	hum	358	19	0.3	2477	9	ADC47904	Adc47904	Human	PRO
286	19	0.3	2477	8	ADA77488	Ada77488	Human	PRO	359	19	0.3	2477	9	ADC04968	Adc04968	Human	sec
287	19	0.3	2477	8	ADBI8228	Adbi8228	CDNA	enco	360	19	0.3	2477	9	ADC79964	Adc79964	Novel	hum
288	19	0.3	2477	8	ADA86911	Ada86911	Novel	hum	361	19	0.3	2477	9	ADC09374	Adc09374	Human	PRO
289	19	0.3	2477	8	ADA16629	Ada16629	Human	sec	362	19	0.3	2477	9	ADC03974	Adc03974	Human	sec
290	19	0.3	2477	8	ADA13058	Ada13058	Human	sec	363	19	0.3	2477	9	ADC03550	Adc03550	Novel	hum
291	19	0.3	2477	8	ADA41926	Ada41926	Human	sec	364	19	0.3	2477	9	ADC03550	Adc03550	Novel	hum
292	19	0.3	2477	8	ADA88014	Ada88014	Novel	hum	365	19	0.3	2477	9	ADC41146	Adc41146	Novel	hum
293	19	0.3	2477	8	ADA46402	Ada46402	Novel	hum	366	19	0.3	2477	9	ADC52285	Adc52285	CDNA	enco
294	19	0.3	2477	8	ADA17273	Ada17273	Human	sec	367	19	0.3	2477	9	ADC53025	Adc53025	CDNA	enco
295	19	0.3	2477	8	ADA43776	Ada43776	Human	sec	368	19	0.3	2477	9	ADC53577	Adc53577	Novel	hum
296	19	0.3	2477	8	ADB28432	Adb28432	CDNA	enco	369	19	0.3	2477	9	ADC51733	Adc51733	CDNA	enco
297	19	0.3	2477	8	ADB28984	Adb28984	CDNA	enco	370	19	0.3	2477	9	ADC02532	Adc02532	Human	PRO
298	19	0.3	2477	8	ADA76936	Ada76936	Human	PRO	371	19	0.3	2477	9	ADC01966	Adc01966	Human	PRO
299	19	0.3	2477	8	ADA88566	Ada88566	Novel	hum	372	19	0.3	2477	9	ADC54148	Adc54148	Human	PRO
300	19	0.3	2477	8	ADA97571	Ada97571	Human	PRO	373	19	0.3	2477	9	ADC92465	Adc92465	Human	PRO
301	19	0.3	2477	8	ADB27328	Adb27328	CDNA	enco	374	19	0.3	2477	9	ADC91361	Adc91361	Human	PRO
302	19	0.3	2477	8	ADB22261	Adb22261	Novel	hum	375	19	0.3	2477	9	ADC03975	Adc03975	Human	PRO
303	19	0.3	2477	8	ACD23630	Acdb23630	Human	PRO	376	19	0.3	2477	9	ADC32272	Adc32272	Novel	hum
304	19	0.3	2477	8	ADA66952	Ada66952	Human	PRO	377	19	0.3	2477	9	ADC22204	Adc22204	CDNA	enco
305	19	0.3	2477	8	ADB22813	Adb22813	Human	PRO	378	19	0.3	2477	9	ADB79428	Adb79428	Human	PRO
306	19	0.3	2477	8	ADB23586	Adb23586	Human	PRO	379	19	0.3	2477	9	ADB41964	Adb41964	Human	PRO
307	19	0.3	2477	8	ADA92308	Ada92308	Novel	hum	380	19	0.3	2477	9	ADB17781	Adb17781	Human	PRO
308	19	0.3	2477	8	ADBI5371	Adbi5371	Human	PRO	381	19	0.3	2477	9	ADB91913	Adb91913	Human	PRO
309	19	0.3	2477	8	ADBI3623	Adbi3623	Novel	hum	382	19	0.3	2477	9	ADB33376	Adb33376	Novel	hum
310	19	0.3	2477	8	ADBI8071	Adbi8071	Novel	hum	383	19	0.3	2477	9	ADB33928	Adb33928	Novel	hum
311	19	0.3	2477	8	ADB66543	Adb66543	Novel	hum	384	19	0.3	2477	9	ADB79980	Adb79980	CDNA	enco
312	19	0.3	2477	8	ADB89623	Adb89623	Human	PRO	385	19	0.3	2477	9	ADD93017	Add93017	Human	PRO
313	19	0.3	2477	8	ADB90355	Adb90355	Human	PRO	386	19	0.3	2477	9	ADB19437	Adb19437	Human	PRO
314	19	0.3	2477	9	ADB77695	Adb77695	Human	sec	387	19	0.3	2477	9	ADB34802	Adb34802	Human	sec
315	19	0.3	2477	9	ADB39456	Adb39456	Novel	hum	388	19	0.3	2477	9	ADB18885	Adb18885	Human	PRO

389	19	0.3	2477	9	ADe43081	462	18	0.3	164	4	ABa37215	AbA37215 Probe #15
390	19	0.3	2477	9	ADe5870	463	18	0.3	164	4	AAK4864	AaK4864 Human bon
391	19	0.3	2477	9	ADe22756	464	18	0.3	164	4	AAK18919	AaK18919 Human bra
392	19	0.3	2477	9	ADd78874	465	18	0.3	164	4	ABe44526	ABe44526 Human liv
393	19	0.3	2477	9	ADe32824	466	18	0.3	164	6	ABe919105	ABe919105 Human gen
394	19	0.3	2477	9	ADe42516	467	18	0.3	209	4	ABL22363	ABl22363 Drosophi
395	19	0.3	2477	9	ADe80532	468	18	0.3	248	4	AAU00188	AAu00188 Human rep
396	19	0.3	2477	9	ADe80532	469	18	0.3	258	7	ABX91685	ABx91685 Murine gen
397	19	0.3	2477	9	ADe80532	470	18	0.3	262	7	AAA42325	AaA42325 Human sec
398	19	0.3	2477	9	ADe80532	471	18	0.3	278	2	AAU22474	AaU22474 Human gen
399	19	0.3	2477	9	ADe80532	472	18	0.3	288	7	ABX87821	ABx87821 Corn ear-
400	19	0.3	2477	10	ADc81068	473	18	0.3	298	4	AAU31739	AAu31739 Novel rat
401	19	0.3	2477	10	ADd76516	474	18	0.3	326	4	AAH50812	AAh50812 Human tum
402	19	0.3	2477	10	ADd76516	475	18	0.3	326	4	AAH50810	AAh50810 Human ben
403	19	0.3	2477	10	ADd76516	476	18	0.3	382	6	ABK64256	ABk64256 Human ORF
404	19	0.3	2477	10	ADd76516	477	18	0.3	384	6	ABN21187	ABn21187 Human ORF
405	19	0.3	2477	10	ADe75732	478	18	0.3	414	4	ABK45931	ABk45931 CDNA enco
406	19	0.3	2477	10	ADe75732	479	18	0.3	420	6	AAK56382	AAk56382 Human enco
407	19	0.3	2477	10	ADe75732	480	18	0.3	429	4	AAI33250	AAi33250 Probe #31
408	19	0.3	2477	10	ADe23308	481	18	0.3	429	4	ABe45948	ABe45948 Human foe
409	19	0.3	2477	10	ADe23308	482	18	0.3	429	4	AAI34602	AAi34602 Probe #32
410	19	0.3	2477	10	ADe23308	483	18	0.3	429	4	ABe44507	ABe44507 Human bre
411	19	0.3	2477	10	ADe23308	484	18	0.3	429	4	ABe44507	ABe44507 Human bre
412	19	0.3	2477	10	ADe23308	485	18	0.3	429	4	ABe44507	ABe44507 Human bre
413	19	0.3	2477	10	ADe23308	486	18	0.3	429	4	ABe44507	ABe44507 Human bre
414	19	0.3	2477	10	ADe23308	487	18	0.3	429	4	ABe44507	ABe44507 Human bre
415	19	0.3	2477	10	ADe23308	488	18	0.3	429	4	ABe44507	ABe44507 Human bre
416	19	0.3	2477	10	ADe23308	489	18	0.3	429	4	ABe44507	ABe44507 Human bre
417	19	0.3	2477	10	ADe23308	490	18	0.3	429	4	ABe44507	ABe44507 Human bre
418	19	0.3	2477	10	ADe23308	491	18	0.3	429	4	ABe44507	ABe44507 Human bre
419	19	0.3	2477	10	ADe23308	492	18	0.3	429	4	ABe44507	ABe44507 Human bre
420	19	0.3	2477	10	ADe23308	493	18	0.3	429	4	ABe44507	ABe44507 Human bre
421	19	0.3	2477	10	ADe23308	494	18	0.3	429	4	ABe44507	ABe44507 Human bre
422	19	0.3	2477	10	ADe23308	495	18	0.3	429	4	ABe44507	ABe44507 Human bre
423	19	0.3	2477	10	ADe23308	496	18	0.3	429	4	ABe44507	ABe44507 Human bre
424	19	0.3	2477	10	ADe23308	497	18	0.3	429	4	ABe44507	ABe44507 Human bre
425	19	0.3	2477	10	ADe23308	498	18	0.3	429	4	ABe44507	ABe44507 Human bre
426	19	0.3	2477	10	ADe23308	499	18	0.3	429	4	ABe44507	ABe44507 Human bre
427	19	0.3	2477	10	ADe23308	500	18	0.3	429	4	ABe44507	ABe44507 Human bre
428	19	0.3	2477	10	ADe23308	501	18	0.3	429	4	ABe44507	ABe44507 Human bre
429	19	0.3	2477	10	ADe23308	502	18	0.3	429	4	ABe44507	ABe44507 Human bre
430	19	0.3	2477	10	ADe23308	503	18	0.3	429	4	ABe44507	ABe44507 Human bre
431	19	0.3	2477	10	ADe23308	504	18	0.3	429	4	ABe44507	ABe44507 Human bre
432	19	0.3	2477	10	ADe23308	505	18	0.3	429	4	ABe44507	ABe44507 Human bre
433	19	0.3	2477	10	ADe23308	506	18	0.3	429	4	ABe44507	ABe44507 Human bre
434	19	0.3	2477	10	ADe23308	507	18	0.3	429	4	ABe44507	ABe44507 Human bre
435	19	0.3	2477	10	ADe23308	508	18	0.3	429	4	ABe44507	ABe44507 Human bre
436	19	0.3	2477	10	ADe23308	509	18	0.3	429	4	ABe44507	ABe44507 Human bre
437	19	0.3	2477	10	ADe23308	510	18	0.3	429	4	ABe44507	ABe44507 Human bre
438	19	0.3	2477	10	ADe23308	511	18	0.3	429	4	ABe44507	ABe44507 Human bre
439	19	0.3	2477	10	ADe23308	512	18	0.3	429	4	ABe44507	ABe44507 Human bre
440	19	0.3	2477	10	ADe23308	513	18	0.3	429	4	ABe44507	ABe44507 Human bre
441	19	0.3	2477	10	ADe23308	514	18	0.3	429	4	ABe44507	ABe44507 Human bre
442	19	0.3	2477	10	ADe23308	515	18	0.3	429	4	ABe44507	ABe44507 Human bre
443	19	0.3	2477	10	ADe23308	516	18	0.3	429	4	ABe44507	ABe44507 Human bre
444	19	0.3	2477	10	ADe23308	517	18	0.3	429	4	ABe44507	ABe44507 Human bre
445	19	0.3	2477	10	ADe23308	518	18	0.3	429	4	ABe44507	ABe44507 Human bre
446	19	0.3	2477	10	ADe23308	519	18	0.3	429	4	ABe44507	ABe44507 Human bre
447	19	0.3	2477	10	ADe23308	520	18	0.3	429	4	ABe44507	ABe44507 Human bre
448	19	0.3	2477	10	ADe23308	521	18	0.3	429	4	ABe44507	ABe44507 Human bre
449	19	0.3	2477	10	ADe23308	522	18	0.3	429	4	ABe44507	ABe44507 Human bre
450	19	0.3	2477	10	ADe23308	523	18	0.3	429	4	ABe44507	ABe44507 Human bre
451	19	0.3	2477	10	ADe23308	524	18	0.3	429	4	ABe44507	ABe44507 Human bre
452	19	0.3	2477	10	ADe23308	525	18	0.3	429	4	ABe44507	ABe44507 Human bre
453	19	0.3	2477	10	ADe23308	526	18	0.3	429	4	ABe44507	ABe44507 Human bre
454	19	0.3	2477	10	ADe23308	527	18	0.3	429	4	ABe44507	ABe44507 Human bre
455	19	0.3	2477	10	ADe23308	528	18	0.3	429	4	ABe44507	ABe44507 Human bre
456	19	0.3	2477	10	ADe23308	529	18	0.3	429	4	ABe44507	ABe44507 Human bre
457	19	0.3	2477	10	ADe23308	530	18	0.3	429	4	ABe44507	ABe44507 Human bre
458	19	0.3	2477	10	ADe23308	531	18	0.3	429	4	ABe44507	ABe44507 Human bre
459	19	0.3	2477	10	ADe23308	532	18	0.3	429	4	ABe44507	ABe44507 Human bre
460	19	0.3	2477	10	ADe23308	533	18	0.3	429	4	ABe44507	ABe44507 Human bre
461	19	0.3	2477	10	ADe23308	534	18	0.3	429	4	ABe44507	ABe44507 Human bre

C 535	18	0.3	729	8	ADB084473	Adb084473	Altioccc	C 608	18	0.3	961	9	ADC17278	Adc17278	CNDA	seqm
C 536	18	0.3	742	9	ADB67959	Adb67959	Human lun	C 609	18	0.3	961	9	ADC14876	Adc14876	Human	PRO
C 537	18	0.3	746	2	AAx40036	AAx40036	Prostate	C 610	18	0.3	961	9	ADC36950	Adc36950	Human	PRO
C 538	18	0.3	759	6	ABK30714	Abk30714	Plant dwa	C 611	18	0.3	961	9	ADC52471	Adc52471	Human	PRO
C 539	18	0.3	759	6	AB149515	Ab149515	Sequence	C 612	18	0.3	961	9	ADC21940	Adc21940	Human	PRO
C 540	18	0.3	797	9	ADB58500	Adb58500	Toxicity-r	C 613	18	0.3	961	9	ADC49971	Adc49971	Human	PRO
C 541	18	0.3	797	9	ADB53098	Adb53098	Primaty-r	C 614	18	0.3	961	9	ADC49170	Adc49170	Human	PRO
C 542	18	0.3	808	4	AAH06598	Aah06598	Human CDN	C 615	18	0.3	961	9	ADC49687	Adc49687	Human	PRO
C 543	18	0.3	814	5	AAH11362	Aah11362	Soybean D	C 616	18	0.3	961	9	ADC47548	Adc47548	Human	PRO
C 544	18	0.3	816	4	AAH69602	Aah69602	Human toe	C 617	18	0.3	961	9	ADC47293	Adc47293	Human	PRO
C 545	18	0.3	816	4	AAH43312	Aah43312	Human liv	C 618	18	0.3	961	9	ADC78168	Adc78168	Human	PRO
C 546	18	0.3	831	9	ADD17516	Add17516	DNA (Seg1	C 619	18	0.3	961	9	ADC06403	Adc06403	Human	PRO
C 547	18	0.3	832	6	AAK09490	Aak09490	Human ova	C 620	18	0.3	961	9	ADC77922	Adc77922	Human	PRO
C 548	18	0.3	833	6	AAK04033	Aak04033	Prostate	C 621	18	0.3	961	9	ADC50885	Adc50885	Human	PRO
C 549	18	0.3	840	6	AAH066625	Aah066625	Isstera	C 622	18	0.3	961	9	ADC51131	Adc51131	Human	PRO
C 550	18	0.3	845	6	AAH98771	Aah98771	Human EST	C 623	18	0.3	961	9	ADC36147	Adc36147	Human	PRO
C 551	18	0.3	849	6	AAK75468	Aak75468	Bacillus	C 624	18	0.3	961	9	ADC50612	Adc50612	Human	PRO
C 552	18	0.3	873	6	AAH67310	Aah67310	Breast sp	C 625	18	0.3	961	9	ADC50366	Adc50366	Human	PRO
C 553	18	0.3	882	5	AAH08650	Aah08650	Mouse can	C 626	18	0.3	961	9	ADC51377	Adc51377	Human	PRO
C 554	18	0.3	899	7	AAH15733	Aah15733	Human can	C 627	18	0.3	961	10	ADC48924	Adc48924	Human	PRO
C 555	18	0.3	921	3	AAH26406	Aah26406	Human sec	C 628	18	0.3	961	10	ADC52281	Adc52281	Human	PRO
C 556	18	0.3	921	5	AAH87191	Aah87191	DNA encod	C 629	18	0.3	961	10	ADC21095	Adc21095	Human	PRO
C 557	18	0.3	921	5	AAH55984	Aah55984	Gene encd	C 630	18	0.3	961	10	ADC05939	Adc05939	Human	PRO
C 558	18	0.3	921	7	AAH39795	Aah39795	Human sec	C 631	18	0.3	961	10	ADC75168	Adc75168	Human	PRO
C 559	18	0.3	921	7	AAH39795	Aah39795	Human sec	C 632	18	0.3	961	10	ADC75914	Adc75914	Human	PRO
C 560	18	0.3	921	7	AAH39795	Aah39795	Human sec	C 633	18	0.3	961	10	ADC86516	Adc86516	Human	PRO
C 561	18	0.3	921	7	AAH39795	Aah39795	Human sec	C 634	18	0.3	961	10	ADC86572	Adc86572	Human	PRO
C 562	18	0.3	961	6	AAH33646	Aah33646	CNDA encd	C 635	18	0.3	961	10	ADC20849	Adc20849	Human	PRO
C 563	18	0.3	961	6	AAH74449	Aah74449	Human CDN	C 636	18	0.3	961	10	ADC839146	Adc839146	Human	PRO
C 564	18	0.3	961	7	ACB66951	Acab66951	CNDA encd	C 637	18	0.3	961	10	ADC05693	Adc05693	Human	PRO
C 565	18	0.3	961	7	ACD68703	Acad68703	Novel hum	C 638	18	0.3	961	10	ADC73678	Adc73678	Human	PRO
C 566	18	0.3	961	7	ACB91235	Acab91235	Novel hum	C 639	18	0.3	961	10	ADC78518	Adc78518	Human	PRO
C 567	18	0.3	961	7	ACD81612	Acad81612	Human CDN	C 640	18	0.3	961	10	ADC21341	Adc21341	Human	PRO
C 568	18	0.3	961	7	ACA60434	Acab60434	Novel hum	C 641	18	0.3	961	10	ADC277456	Adc277456	Human	PRO
C 569	18	0.3	961	7	ACA58881	Acab58881	CNDA encd	C 642	18	0.3	961	10	ADC20603	Adc20603	Human	PRO
C 570	18	0.3	961	7	ACA64057	Acab64057	CNDA encd	C 643	18	0.3	961	10	ADC75668	Adc75668	Human	PRO
C 571	18	0.3	961	7	ACA91321	Acab91321	CNDA encd	C 644	18	0.3	961	10	ADC74184	Adc74184	Human	PRO
C 572	18	0.3	961	7	ACD45220	Acad45220	Human sec	C 645	18	0.3	961	10	ADC74430	Adc74430	Human	PRO
C 573	18	0.3	961	7	ACA93768	Acab93768	Human CDN	C 646	18	0.3	961	10	ADC76160	Adc76160	Human	PRO
C 574	18	0.3	961	7	ACA67342	Acab67342	CNDA encd	C 647	18	0.3	961	10	ADC85652	Adc85652	Novel	hum
C 575	18	0.3	961	7	ACA68607	Acab68607	Novel hum	C 648	18	0.3	961	10	ADC805201	Adc805201	Human	PRO
C 576	18	0.3	961	7	ACA66315	Acab66315	Novel hum	C 649	18	0.3	961	10	ADC75414	Adc75414	Human	PRO
C 577	18	0.3	961	7	ACD02369	Acad02369	Novel hum	C 650	18	0.3	961	10	ADC76958	Adc76958	Novel	hum
C 578	18	0.3	961	7	ACA89360	Acab89360	Novel hum	C 651	18	0.3	961	10	ADC86726	Adc86726	Novel	hum
C 579	18	0.3	961	7	ACA68997	Acab68997	Novel hum	C 652	18	0.3	961	10	ADC78194	Adc78194	Novel	hum
C 580	18	0.3	961	7	ACA98519	Acab98519	Human PRO	C 653	18	0.3	961	10	ADC77702	Adc77702	Novel	hum
C 581	18	0.3	961	8	ACA63444	Acab63444	CNDA encd	C 654	18	0.3	961	10	ADC77948	Adc77948	Novel	hum
C 582	18	0.3	961	8	ABT44336	Abt44336	Human PRO	C 655	18	0.3	961	10	ADC85406	Adc85406	Novel	hum
C 583	18	0.3	961	8	ADB17200	Adb17200	Human CDN	C 656	18	0.3	961	10	ADC73938	Adc73938	Human	PRO
C 584	18	0.3	961	8	ACH03647	Aach03647	Human sec	C 657	18	0.3	961	10	ADC74676	Adc74676	Human	PRO
C 585	18	0.3	961	8	ADA20005	Ada20005	Novel hum	C 658	18	0.3	961	10	ADC77204	Adc77204	Novel	hum
C 586	18	0.3	961	8	ADA17388	Ada17388	Human CDN	C 659	18	0.3	961	10	ADC85898	Adc85898	Novel	hum
C 587	18	0.3	961	8	ADA20177	Ada20177	Novel hum	C 660	18	0.3	961	10	ADC75457	Adc75457	Human	PRO
C 588	18	0.3	961	8	ACD82161	Acad82161	Human sec	C 661	18	0.3	961	10	ADC74822	Adc74822	Human	PRO
C 589	18	0.3	961	8	ABT44619	Abt44619	Human PRO	C 662	18	0.3	961	10	AAH53900	Aah53900	Human	PRO
C 590	18	0.3	961	8	ACD82286	Acad82286	Human sec	C 663	18	0.3	966	7	ABZ51974	Abz51974	Human	PRO
C 591	18	0.3	961	8	ADA00474	Ada00474	Human sec	C 664	18	0.3	966	7	ADC03553	Adc03553	Human	PRO
C 592	18	0.3	961	8	ABT43992	Abt43992	Human mem	C 665	18	0.3	1000	9	ADA03013	Ada03013	Mouse Ncf	
C 593	18	0.3	961	8	ADB85716	Adb85716	Novel hum	C 666	18	0.3	1020	8	ADB27271	Adb27271	Mouse Ncf	
C 594	18	0.3	961	8	ADB85716	Adb85716	Novel hum	C 667	18	0.3	1020	9	ADC85493	Adc85493	Mouse Ncf	
C 595	18	0.3	961	8	ADB80817	Adb80817	Novel hum	C 668	18	0.3	1023	8	ADA48577	Ada48577	Rice gene	
C 596	18	0.3	961	8	ADB80817	Adb80817	Novel hum	C 669	18	0.3	1036	7	ABT32209	Abt32209	Human PRO	
C 597	18	0.3	961	8	ADB78440	Adb78440	Novel hum	C 670	18	0.3	1040	7	AAQ29530	Aaq29530	Pancreas	
C 598	18	0.3	961	8	ADB85088	Adb85088	Human PRO	C 671	18	0.3	1051	6	AAH58581	Aah58581	Human PRO	
C 599	18	0.3	961	9	ADB78194	Adb78194	Novel hum	C 672	18	0.3	1084	4	AAH64040	Aah64040	Human PRO	
C 600	18	0.3	961	9	ADB887260	Adb887260	Human PRO	C 673	18	0.3	1090	2	AAH64040	Aah64040	Human PRO	
C 601	18	0.3	961	9	ADB884842	Adb884842	Human PRO	C 674	18	0.3	1113	4	ABT46946	Abt46946	Prokaryot	
C 602	18	0.3	961	9	ADB88395	Adb88395	Human PRO	C 675	18	0.3	1119	7	ACA46946	Acab46946	Prokaryot	
C 603	18	0.3	961	9	ADB868202	Adb868202	Human PRO	C 676	18	0.3	1129	7	ABZ80037	Abz80037	Human AGE	
C 604	18	0.3	961	9	ADB83957	Adb83957	Novel hum	C 677	18	0.3	1131	2	AAH89663	Aah89663	Human AGE	
C 605	18	0.3	961	9	ADB73112	Adb73112	Novel hum	C 678	18	0.3	1131	2	AAQ00029	Aaq00029	Sequence	
C 606	18	0.3	961	9	ADB91019	Adb91019	Novel hum	C 679	18	0.3	1161	6	ABN1470	Abn1470	Staphyloc	
C 607	18	0.3	961	9	ADC07099	Adc07099	Human PRO	C 680	18	0.3	1214	4	AAH24443	Aah24443	Human Bre	

C 681	18	0.3	1236	4	AAK78770	AAK78770	Human imm	C 754	18	0.3	2724	4	ABLI6706	ABLI6706	Drosophila
C 682	18	0.3	1236	4	AAK78769	AAK78769	Human imm	C 755	18	0.3	2810	7	ABZ24708	ABZ24708	Human cel
C 683	18	0.3	1242	5	AAK87318	AAK87318	DNA encod	C 756	18	0.3	2888	4	ABLI6825	ABLI6825	Drosophila
C 684	18	0.3	1242	6	AAK93932	AAK93932	DNA of th	C 757	18	0.3	2940	2	AAK60397	AAK60397	Human pol
C 685	18	0.3	1317	7	ACA36042	ACA36042	Prokaryot	C 758	18	0.3	3000	2	AAK84399	AAK84399	S. capsul
C 686	18	0.3	1317	7	ACA31646	ACA31646	Prokaryot	C 759	18	0.3	3044	4	ABLI06599	ABLI06599	Drosophila
C 687	18	0.3	1335	4	AAK61707	AAK61707	Human imm	C 760	18	0.3	3093	6	ABZ13765	ABZ13765	Arabidops
C 688	18	0.3	1362	7	ACA30833	ACA30833	Prokaryot	C 761	18	0.3	3183	7	AAI52258	AAI52258	184P182-r
C 689	18	0.3	1385	4	AAH99262	AAH99262	Human pro	C 762	18	0.3	3183	7	AAI52259	AAI52259	184P182 g
C 690	18	0.3	1395	8	AAH2712	AAH2712	Alloioococ	C 763	18	0.3	3183	7	AAI52262	AAI52262	184P182-r
C 691	18	0.3	1405	6	AAI48815	AAI48815	Murine an	C 764	18	0.3	3183	7	AAI52268	AAI52268	184P182-g
C 692	18	0.3	1406	7	ABX05078	ABX05078	Human nov	C 765	18	0.3	3183	7	AAI52260	AAI52260	184P182 g
C 693	18	0.3	1425	8	ADA03012	ADA03012	Mouse Ncf	C 766	18	0.3	3183	7	AAI52265	AAI52265	184P182 g
C 694	18	0.3	1425	9	ADB72750	ADB72750	Mouse Ncf	C 767	18	0.3	3183	7	AAI52267	AAI52267	184P182 g
C 695	18	0.3	1425	9	ADB72750	ADB72750	Mouse Ncf	C 768	18	0.3	3183	7	AAI52264	AAI52264	184P182 g
C 696	18	0.3	1425	9	ADB72750	ADB72750	Mouse Ncf	C 769	18	0.3	3183	7	AAI52261	AAI52261	184P182 g
C 697	18	0.3	1427	7	ACD07381	ACD07381	Rice Myd-	C 770	18	0.3	3183	7	AAI52266	AAI52266	184P182 g
C 698	18	0.3	1460	2	AAI38075	AAI38075	Sequence	C 771	18	0.3	3183	7	AAI52261	AAI52261	184P182 g
C 699	18	0.3	1461	8	ACD26032	ACD26032	Human nov	C 772	18	0.3	3227	4	AAK80792	AAK80792	184P182 g
C 700	18	0.3	1507	3	AD30526	AD30526	Human nov	C 773	18	0.3	3227	4	AAK80791	AAK80791	184P182 g
C 701	18	0.3	1520	3	ABO70450	ABO70450	Wild-type	C 774	18	0.3	3227	4	AAK80791	AAK80791	184P182 g
C 702	18	0.3	1520	6	AB542341	AB542341	Fruit fly	C 775	18	0.3	3251	4	ABLI13875	ABLI13875	Drosophila
C 703	18	0.3	1520	6	AB54411	AB54411	Fruit fly	C 776	18	0.3	3347	4	ABLI13875	ABLI13875	Drosophila
C 704	18	0.3	1554	7	ACF70169	ACF70169	Phototrab	C 777	18	0.3	3347	4	ABLI13875	ABLI13875	Drosophila
C 705	18	0.3	1559	4	ABLI2362	ABLI2362	Drosophila	C 778	18	0.3	3366	6	ABT07269	ABT07269	Human CDG
C 706	18	0.3	1581	5	AAH21791	AAH21791	Mouse G4S	C 779	18	0.3	3408	9	ADB69555	ADB69555	C. neofor
C 707	18	0.3	1604	9	ADB59144	ADB59144	Toxicity-	C 780	18	0.3	3437	2	AAZ07545	AAZ07545	Human vit
C 708	18	0.3	1607	3	AAK54954	AAK54954	Arabidops	C 781	18	0.3	3510	2	AAZ07545	AAZ07545	Human vit
C 709	18	0.3	1609	3	AAK53307	AAK53307	Arabidops	C 782	18	0.3	3622	1	AAH60204	AAH60204	Drosophila
C 710	18	0.3	1617	4	AAH12576	AAH12576	Human pro	C 783	18	0.3	3659	4	ABLI22034	ABLI22034	Drosophila
C 711	18	0.3	1633	7	ABQ77395	ABQ77395	Human RNA	C 784	18	0.3	3688	4	ADB79946	ADB79946	Human kin
C 712	18	0.3	1648	6	ABQ60809	ABQ60809	Unidentif	C 785	18	0.3	3854	6	AAH17848	AAH17848	Human nuc
C 713	18	0.3	1751	4	AAI58611	AAI58611	Human pol	C 786	18	0.3	3985	4	ACF35658	ACF35658	Human kin
C 714	18	0.3	1751	8	ADB48589	ADB48589	Novel hum	C 787	18	0.3	4077	8	AAH53178	AAH53178	Novel hum
C 715	18	0.3	1771	8	ADB48589	ADB48589	Novel hum	C 788	18	0.3	4189	7	ABX70939	ABX70939	Novel hum
C 716	18	0.3	1771	8	ADB48589	ADB48589	Novel hum	C 789	18	0.3	4202	7	ABLI17590	ABLI17590	Drosophila
C 717	18	0.3	1771	8	ADB48589	ADB48589	Novel hum	C 790	18	0.3	4233	4	AAK53303	AAK53303	Human pol
C 718	18	0.3	1771	8	ADB48589	ADB48589	Novel hum	C 791	18	0.3	4287	4	AAK52319	AAK52319	Human pol
C 719	18	0.3	1827	3	AAFI8251	AAFI8251	Lung cano	C 792	18	0.3	4304	4	AAK52319	AAK52319	Human pol
C 720	18	0.3	1840	3	AAK94394	AAK94394	Human ful	C 793	18	0.3	4408	7	AAI02765	AAI02765	Drosophila
C 721	18	0.3	1882	7	ABT41827	ABT41827	Rat OCT-1	C 794	18	0.3	4461	9	ADB99854	ADB99854	Drosophila
C 722	18	0.3	1882	7	ABT41827	ABT41827	Toxicity-	C 795	18	0.3	4473	9	ADB99854	ADB99854	Drosophila
C 723	18	0.3	1882	9	ADB59275	ADB59275	Toxicity-	C 796	18	0.3	4594	5	AAV03326	AAV03326	DNA encod
C 724	18	0.3	1882	9	ADB59275	ADB59275	Toxicity-	C 797	18	0.3	4596	2	AAV03326	AAV03326	DNA encod
C 725	18	0.3	1926	2	AAV82522	AAV82522	Primary r	C 798	18	0.3	4596	5	AAV03326	AAV03326	DNA encod
C 726	18	0.3	1948	4	ABAS7042	ABAS7042	Human fce	C 799	18	0.3	4596	6	AAK51171	AAK51171	Drosophila
C 727	18	0.3	1948	4	ABAS7042	ABAS7042	Human fce	C 800	18	0.3	4596	6	AAK51171	AAK51171	Drosophila
C 728	18	0.3	1959	5	AAAF6317	AAAF6317	Pseudomon	C 801	18	0.3	4596	5	AAK51171	AAK51171	Drosophila
C 729	18	0.3	2000	4	ABLI20725	ABLI20725	Drosophila	C 802	18	0.3	4599	5	AAK51171	AAK51171	Drosophila
C 730	18	0.3	2006	4	ADC32458	ADC32458	Human nov	C 803	18	0.3	4604	2	AAV41327	AAV41327	DNA encod
C 731	18	0.3	2014	3	ACA23442	ACA23442	Human sec	C 804	18	0.3	4604	6	AAK51171	AAK51171	Drosophila
C 732	18	0.3	2084	7	ACA23442	ACA23442	Human sec	C 805	18	0.3	4604	6	AAK51171	AAK51171	Drosophila
C 733	18	0.3	2213	4	ABLI03833	ABLI03833	Drosophila	C 806	18	0.3	4604	6	AAK51171	AAK51171	Drosophila
C 734	18	0.3	2233	4	ABLI17591	ABLI17591	Drosophila	C 807	18	0.3	4604	6	AAK51171	AAK51171	Drosophila
C 735	18	0.3	2253	6	ABZ14355	ABZ14355	Arabidops	C 808	18	0.3	4673	2	AAI12787	AAI12787	Drosophila
C 736	18	0.3	2321	4	ABLI20727	ABLI20727	Drosophila	C 809	18	0.3	4727	2	ABLI13386	ABLI13386	Drosophila
C 737	18	0.3	2335	7	ACC51207	ACC51207	Human pik	C 810	18	0.3	4794	4	ABLI19928	ABLI19928	Drosophila
C 738	18	0.3	2346	4	ABLI22154	ABLI22154	Drosophila	C 811	18	0.3	4847	2	AAV58943	AAV58943	Drosophila
C 739	18	0.3	2346	4	ABLI22154	ABLI22154	Drosophila	C 812	18	0.3	4847	2	AAV58943	AAV58943	Drosophila
C 740	18	0.3	2379	6	ABZ14366	ABZ14366	Arabidops	C 813	18	0.3	5019	4	AAI04636	AAI04636	Human tes
C 741	18	0.3	2414	6	ABLI22150	ABLI22150	Drosophila	C 814	18	0.3	5049	4	ABLI97603	ABLI97603	Human tes
C 742	18	0.3	2415	6	ABK73868	ABK73868	Bacillus	C 815	18	0.3	5049	4	ABLI97603	ABLI97603	Human tes
C 743	18	0.3	2430	5	AAK90898	AAK90898	DNA encod	C 816	18	0.3	5135	4	ABLI03464	ABLI03464	Drosophila
C 744	18	0.3	2484	4	ABLI03465	ABLI03465	Drosophila	C 817	18	0.3	5168	6	AAK594716	AAK594716	Drosophila
C 745	18	0.3	2489	4	ABLI22152	ABLI22152	Drosophila	C 818	18	0.3	5211	5	AAK594716	AAK594716	Drosophila
C 746	18	0.3	2505	2	AAK69865	AAK69865	Human gro	C 819	18	0.3	5289	9	ADB69194	ADB69194	Drosophila
C 747	18	0.3	2510	6	AAK83456	AAK83456	Human nov	C 820	18	0.3	5350	4	ABLI06598	ABLI06598	Drosophila
C 748	18	0.3	2544	9	ADC30652	ADC30652	Human nov	C 821	18	0.3	5359	9	ADCO1427	ADCO1427	Drosophila
C 749	18	0.3	2596	4	AAI86538	AAI86538	Human pol	C 822	18	0.3	5323	9	ADB53664	ADB53664	Human pro
C 750	18	0.3	2602	4	ABLI26431	ABLI26431	Drosophila	C 823	18	0.3	5433	9	ADB53663	ADB53663	Human pro
C 751	18	0.3	2613	6	ABK28647	ABK28647	Human CDN	C 824	18	0.3	5484	4	ABLI13874	ABLI13874	Drosophila
C 752	18	0.3	2626	6	ADB63577	ADB63577	Human CDN	C 825	18	0.3	5557	5	AAH21794	AAH21794	Drosophila
C 753	18	0.3	2670	9	ADB69916	ADB69916	C. neofor	C 826	18	0.3	5699	5	AAK89500	AAK89500	DNA encod

827	18	0.3	5721	4	ABLI1973	AbI11973 Drosophila	900	18	0.3	110000	6	ABA03041_03	Continuation (4 of
C 828	18	0.3	5938	7	ABK84688	AbK46888 Human CDN	C 901	18	0.3	110000	7	ACF65385_3	Continuation (4 of
C 829	18	0.3	6186	7	ABT19929	AbT19929 Aspergillus	C 902	18	0.3	110000	7	ACF67367_32	Continuation (33 of
C 830	18	0.3	6196	7	ABT18115	AbT18115 Aspergillus	C 903	18	0.3	110000	7	ABD12064_06	Continuation (7 of
C 831	18	0.3	6242	9	ADD71201	AbD71201 Human int	C 904	18	0.3	110000	8	ABD12064_15	Continuation (16 of
C 832	18	0.3	6601	7	ABT19996	AbT19996 Aspergillus	C 905	18	0.3	122186	4	AAH89560	AAH89560 Human his
C 833	18	0.3	6601	7	ABT18182	AbT18182 Aspergillus	C 906	18	0.3	123219	4	AAH88703	AAH88703 Human DNA
C 834	18	0.3	6679	4	ABLI20726	AbI20726 Drosophila	C 907	18	0.3	165199	6	ABK83460	ABK83460 Human CDN
C 835	18	0.3	6690	4	ABLI6769	AbI6769 Drosophila	C 908	18	0.3	188888	6	ABQ75562	ABQ75562 Human rel
C 836	18	0.3	7033	3	AAZ36989	AAZ36989 Human per	C 909	18	0.3	197997	7	AAI54074	AAI54074 Human tra
C 837	18	0.3	7060	4	ABLI20724	AbI20724 Drosophila	C 910	18	0.3	198161	6	AAK83564	AAK83564 Human CDN
C 838	18	0.3	7079	4	ABLI03832	AbI03832 Drosophila	C 911	18	0.3	220895	6	ABK84798	ABK84798 Human CDN
C 839	18	0.3	7228	4	AAI04695	AAI04695 Human rep	C 912	18	0.3	349980	6	ABQ81847	ABQ81847 Human h1b
C 840	18	0.3	7228	4	ABLI97602	AbI97602 Human tes	C 913	17	0.3	20	6	AAI35758	AAI35758 Human h1b
C 841	18	0.3	7330	4	ABLI06499	AbI06499 Drosophila	C 914	17	0.3	22	7	ADA06040	ADA06040 Human NOV
C 842	18	0.3	7568	5	AAH80823	AAH80823 DNA encod	C 915	17	0.3	22	7	ABT33573	ABT33573 NOV probe
C 843	18	0.3	8097	4	ABLI1972	AbI1972 Drosophila	C 916	17	0.3	22	7	ABT06340	ABT06340 Human NOV
C 844	18	0.3	8254	5	ABAI1475	ABAI1475 Human ner	C 917	17	0.3	25	6	ABT06367	ABT06367 Human NOV
C 845	18	0.3	9220	5	ABV24414	ABV24414 Human pro	C 918	17	0.3	25	6	ABT06370	ABT06370 Human NOV
C 846	18	0.3	9220	5	ABV25265	ABV25265 Human pro	C 919	17	0.3	25	6	ABT06343	ABT06343 Human NOV
C 847	18	0.3	9416	6	ABK83800	AbK83800 Human CDN	C 920	17	0.3	32	2	AAI42078	AAI42078 Human ery
C 848	18	0.3	9416	7	ACF34559	ACF34559 Gene enco	C 921	17	0.3	54	2	AAQ42517	AAQ42517 Sequence
C 849	18	0.3	10543	4	ABLI06498	AbI06498 Drosophila	C 922	17	0.3	60	6	ABH42958	ABH42958 Human spl
C 850	18	0.3	11304	4	AAI03113	AAI03113 Human rep	C 923	17	0.3	100	7	ACD76585	ACD76585 E. coli K
C 851	18	0.3	11950	9	ADBS3871	ADBS3871 Human pro	C 924	17	0.3	116	3	AAI42517	AAI42517 Sequence
C 852	18	0.3	13993	9	ADBS31304	ADBS31304 Testeter	C 925	17	0.3	117	4	ABH74494	ABH74494 Human sec
C 853	18	0.3	14070	2	AAZ09525	AAZ09525 Human Apo	C 926	17	0.3	117	4	AAI54965	AAI54965 Human bra
C 854	18	0.3	14070	7	ACA64742	ACA64742 Apolipop	C 927	17	0.3	117	4	AAK49135	AAK49135 Human bra
C 855	18	0.3	14121	6	ABT13015	AbT13015 Human apo	C 928	17	0.3	117	4	AAK22962	AAK22962 Human bra
C 856	18	0.3	14121	7	ACG62114	ACG62114 Human ali	C 929	17	0.3	117	4	AAK22962	AAK22962 Human bra
C 857	18	0.3	15254	4	AAK83135	AAK83135 Human imm	C 930	17	0.3	135	4	AAH75398	AAH75398 Human col
C 858	18	0.3	15254	4	AAK83137	AAK83137 Human imm	C 931	17	0.3	138	4	AAI49333	AAI49333 Human bra
C 859	18	0.3	15255	4	AAK83136	AAK83136 Human imm	C 932	17	0.3	138	4	ABH51150	ABH51150 Human bra
C 860	18	0.3	15256	4	AAK83134	AAK83134 Human imm	C 933	17	0.3	138	4	AAK43360	AAK43360 Human bra
C 861	18	0.3	22873	9	ADCB7198	ADCB7198 Human GPC	C 934	17	0.3	138	4	AAK17456	AAK17456 Human bra
C 862	18	0.3	24259	9	AAH46691	AAH46691 Tumour su	C 935	17	0.3	138	5	AAI09615	AAI09615 Human bra
C 863	18	0.3	28315	9	ADCB6254	ADCB6254 Human GPC	C 936	17	0.3	162	2	ABK14569	ABK14569 Human pap
C 864	18	0.3	30610	5	ABAI5643	ABAI5643 Human ner	C 937	17	0.3	193	6	ABK14569	ABK14569 Human pap
C 865	18	0.3	31871	8	ADCB6672	ADCB6672 Connectiv	C 938	17	0.3	211	7	ABH17523	ABH17523 Human ner
C 866	18	0.3	32200	8	ADBS42917	ADBS42917 Connectiv	C 939	17	0.3	215	5	ABH17522	ABH17522 Human ner
C 867	18	0.3	32200	8	ADBS42917	ADBS42917 Connectiv	C 940	17	0.3	215	5	ABH17522	ABH17522 Human ner
C 868	18	0.3	34269	4	AAK85168	AAK85168 Human imm	C 941	17	0.3	222	5	AAI52284	AAI52284 Human AFP
C 869	18	0.3	34269	4	AAK85168	AAK85168 Human imm	C 942	17	0.3	222	5	AAI52284	AAI52284 Human AFP
C 870	18	0.3	36776	4	AAK85221	AAK85221 Human imm	C 943	17	0.3	225	5	AAI52284	AAI52284 Human AFP
C 871	18	0.3	36776	4	AAK85221	AAK85221 Human imm	C 944	17	0.3	225	5	AAI52284	AAI52284 Human AFP
C 872	18	0.3	36643	8	ADH72749	ADH72749 Mouse NCF	C 945	17	0.3	236	4	ABH68827	ABH68827 Human bra
C 873	18	0.3	36643	9	ADH72749	ADH72749 Mouse NCF	C 946	17	0.3	236	4	ABH68827	ABH68827 Human bra
C 874	18	0.3	36643	9	ADH72749	ADH72749 Mouse NCF	C 947	17	0.3	236	4	ABH68827	ABH68827 Human bra
C 875	18	0.3	40328	4	AAZ92584	AAZ92584 Human DAZ	C 948	17	0.3	236	4	ABH68827	ABH68827 Human bra
C 876	18	0.3	40328	4	AAZ92584	AAZ92584 Human DAZ	C 949	17	0.3	236	4	ABH68827	ABH68827 Human bra
C 877	18	0.3	40742	4	AAK79886	AAK79886 Human imm	C 950	17	0.3	236	4	ABH68827	ABH68827 Human bra
C 878	18	0.3	40742	4	AAK79886	AAK79886 Human imm	C 951	17	0.3	236	4	ABH68827	ABH68827 Human bra
C 879	18	0.3	43795	3	AAZ92583	AAZ92583 Human DAZ	C 952	17	0.3	236	4	ABH68827	ABH68827 Human bra
C 880	18	0.3	44979	9	ADCB9873	ADCB9873 Partial g	C 953	17	0.3	236	5	AAI08987	AAI08987 Human bra
C 881	18	0.3	45000	9	ADCB9873	ADCB9873 Partial g	C 954	17	0.3	236	5	AAI08987	AAI08987 Human bra
C 882	18	0.3	47670	4	ABLI16824	AbI16824 Drosophila	C 955	17	0.3	236	6	ABH16418	ABH16418 Human bra
C 883	18	0.3	50368	4	ABLI16768	AbI16768 Drosophila	C 956	17	0.3	236	6	ABH16418	ABH16418 Human bra
C 884	18	0.3	53522	6	AAAD30228	AAAD30228 Human PKD	C 957	17	0.3	242	4	AAK63336	AAK63336 Human bra
C 885	18	0.3	53522	6	AAAD30228	AAAD30228 Human PKD	C 958	17	0.3	249	2	AAK52018	AAK52018 Human bra
C 886	18	0.3	53522	6	AAAD30228	AAAD30228 Human PKD	C 959	17	0.3	258	3	AAI10218	AAI10218 Human bra
C 887	18	0.3	53577	2	AAI19851	AAI19851 Human POL	C 960	17	0.3	261	8	ACH44287	ACH44287 Human bra
C 888	18	0.3	53577	2	AAI19851	AAI19851 Human POL	C 961	17	0.3	273	4	ABH62999	ABH62999 Human bra
C 889	18	0.3	67571	7	AAAD53224_5	AAAD53224_5 Human POL	C 962	17	0.3	273	4	ABH62999	ABH62999 Human bra
C 890	18	0.3	72149	9	ADCB85476	ADCB85476 Human Mst	C 963	17	0.3	278	7	ABH7360	ABH7360 Human bra
C 891	18	0.3	76595	8	ADA02996	ADA02996 Human Mst	C 964	17	0.3	284	4	ABH7360	ABH7360 Human bra
C 892	18	0.3	96595	8	ADA02996	ADA02996 Human Mst	C 965	17	0.3	303	7	ACD97447	ACD97447 Human bra
C 893	18	0.3	110000	2	AAI991990_00	AAI991990_00 Human Mst	C 966	17	0.3	303	7	ACD97447	ACD97447 Human bra
C 894	18	0.3	110000	4	AAI99682_10	AAI99682_10 Human Mst	C 967	17	0.3	303	7	ACD97447	ACD97447 Human bra
C 895	18	0.3	110000	4	AAI99682_11	AAI99682_11 Human Mst	C 968	17	0.3	303	7	ACD97447	ACD97447 Human bra
C 896	18	0.3	110000	4	AAI99682_12	AAI99682_12 Human Mst	C 969	17	0.3	303	7	ACD97447	ACD97447 Human bra
C 897	18	0.3	110000	4	AAI99682_13	AAI99682_13 Human Mst	C 970	17	0.3	315	5	ABV01349	ABV01349 Human bra
C 898	18	0.3	110000	4	AAI99682_14	AAI99682_14 Human Mst	C 971	17	0.3	321	4	ABH74039	ABH74039 Human bra
C 899	18	0.3	110000	4	ABO69245_03	ABO69245_03 Human Mst	C 972	17	0.3	321	4	AAI54489	AAI54489 Human bra
C 900	18	0.3	110000	6	ABO6797_02	ABO6797_02 Human Mst	C 973	17	0.3	321	4	AAI54489	AAI54489 Human bra

973	17	0.3	321	4	AAK22493	AAK22493	Human	b2a
974	17	0.3	321	4	ABG48341	ABG48341	Human	l1v
C 975	17	0.3	321	6	ABN94568	ABN94568	Gene	#106
C 976	17	0.3	326	6	ABN21344	ABN21344	Human	ORF
C 977	17	0.3	338	6	ABN13835	ABN13835	Human	ner
C 978	17	0.3	345	6	ABV96346	ABV96346	Human	pan
C 979	17	0.3	346	4	AAK55956	AAK55956	Human	imm
C 980	17	0.3	351	7	ABX52688	ABX52688	Bovine	ES
C 981	17	0.3	351	7	AAV86902	AAV86902	EST	clone
C 982	17	0.3	351	7	ACC44893	ACC44893	Mouse	TSH
C 983	17	0.3	354	8	ACD24576	ACD24576	DNA	clone
C 984	17	0.3	356	4	AAAD3894	AAAD3894	Human	tar
C 985	17	0.3	363	4	AAK92899	AAK92899	Human	CDN
C 986	17	0.3	366	4	AAK92899	AAK92899	Human	sec
C 987	17	0.3	371	6	ABV95176	ABV95176	Human	pan
C 988	17	0.3	372	6	ACH47740	ACH47740	Human	inf
C 989	17	0.3	373	6	ABV95612	ABV95612	Human	pan
C 990	17	0.3	374	6	ABV95993	ABV95993	Human	pan
C 991	17	0.3	377	8	ACL24578	ACL24578	DNA	clone
C 992	17	0.3	378	4	AAI13465	AAI13465	Probe	#33
C 993	17	0.3	378	4	AAI13481	AAI13481	Human	foe
C 994	17	0.3	378	4	ABAB5162	ABAB5162	Human	bre
C 995	17	0.3	378	4	ABA44706	ABA44706	Human	b2e
C 996	17	0.3	378	4	ABA24907	ABA24907	Probe	#13
C 997	17	0.3	378	4	AAK28874	AAK28874	Human	don
C 998	17	0.3	378	4	AAK03423	AAK03423	Human	b2a
C 999	17	0.3	378	4	ABE28491	ABE28491	Human	l1v
C1000	17	0.3	378	5	AAI03344	AAI03344	Probe	#33

ALIGNMENTS

RESULT 1

ABS63437

ID ABS63437 standard; cDNA; 5691 BP.

AC ABS63437;

DT 15-NOV-2002 (first entry)

XX Human cDNA, homologous to plexin, designated NOVA.

Human; gene; ss; NOV; NOVX; NOVX-associated disorder; cardiomyopathy;
 atherosclerosis; diabetes; cell signalling; metabolic pathway;
 cellular receptor; downstream effector; cancer; gene therapy;
 hypertension; congenital heart defect; aortic stenosis; obesity;
 infectious disease; anorexia; Alzheimer's disease; Parkinson's disease;
 neurodegenerative disorder; haemophilia; dyslipidemia; vaccine;
 haematopoietic disease; scleroderma; fertility; immunogen;
 idiopathic thrombocytopenic purpura; graft versus host disease;
 Crohn's disease; multiple sclerosis; cirrhosis; autoimmune disease;
 systemic lupus erythematosus; asthma; arthritis; psoriasis; allergy;
 stroke; anxiety; Lesch-Nyhan syndrome; schizophrenia; cerebellar ataxia;
 pain; alcoholism; transgenic.

XX Homo sapiens.

XX MO200226826-A2.

XX 04-APR-2002.

XX 27-SEP-2001; 2001WO-US042336.

XX 27-SEP-2000; 2000US-0235631P.
 PR 27-SEP-2000; 2000US-0235633P.
 PR 27-SEP-2000; 2000US-0235808P.
 PR 27-SEP-2000; 2000US-0236064P.
 PR 27-SEP-2000; 2000US-0236065P.
 PR 27-SEP-2000; 2000US-0236066P.
 PR 28-SEP-2000; 2000US-0236135P.
 PR 03-OCT-2000; 2000US-0237434P.
 PR 05-OCT-2000; 2000US-0238321P.

PR 06-OCT-2000; 2000US-0238396P.
 PR 06-OCT-2000; 2000US-0238399P.
 PR 16-MAR-2001; 2001US-0276667P.
 PR 31-MAY-2001; 2001US-0294823P.
 PR 12-JUL-2001; 2001US-0304868P.
 PR 26-SEP-2001; 2001US-00235631.

XX (CURA-) CURAGEN CORP.

XX Gerlach VL, MacDougall JR, Smithson G, Millet I, Stone D;

PI Gunther E, Ellerman K, Grosse WM, Alsobrook JP, Lepley DM;

PI Burgess CE, Padigaru M, Kekuda R, Spytek KA, Leach MD, Shinkets RA;

XX WPI; 2002-499860/53.

XX P-PSDB; ABG78364.

XX Novel isolated NOVX polypeptides and polynucleotides homologous to

XX attractin, plexin, papin-like family of proteins, useful for treating

XX atherosclerosis, diabetes, cancer, Alzheimer's disease, hemophilia and

XX stroke.

XX Claim 8; Page 58-59; 308pp; English.

XX The invention discloses the isolated human polypeptides, and

XX polynucleotides encoding them, that have been designated NOVX. The

XX polypeptides, polynucleotides and antibodies are useful in treating or

XX preventing a NOVX-associated disorder which is cardiomyopathy;

XX cell signal processing and metabolic pathway modulation. They can also be

XX used in determining the presence of, or predisposition to, a disease

XX associated with altered levels of the polypeptides and polynucleotides of

XX any one of the 13 sequences (NOVX-NOV8), for raising antibodies, for

XX identifying an agent that binds to, or that modulates the expression or

XX activity of the polypeptide, for identifying an agent which is cellular

XX receptor or downstream effector, for treating or preventing a NOVX-

XX associated disorder and as a pharmaceutical composition comprising the

XX polypeptide, polynucleotide or the antibody. The polypeptides and

XX polynucleotides are useful in diagnostic applications (e.g. as a marker

XX for cancerous cells or tissue types) where their amounts are assessed, or

XX for the manufacture of a medicament (e.g. gene therapy) for treating or

XX preventing disorders or syndromes such as hypertension, congenital heart

XX defects, aortic stenosis, obesity, infectious diseases, anorexia, cancer,

XX Alzheimer's disease, Parkinson's disorders, neurodegenerative disorders,

XX haemophilia, dyslipidemia, haematopoietic diseases, scleroderma,

XX fertility, idiopathic thrombocytopenic purpura, graft versus host

XX diseases, Crohn's disease, multiple sclerosis, cirrhosis, autoimmune

XX disease, systemic lupus erythematosus, asthma, arthritis, psoriasis,

XX allergy, stroke, anxiety, Lesch-Nyhan syndrome, schizophrenia, cerebellar

XX ataxia, pain and alcoholism. They may also be used as immunogens to

XX transgenic cells containing a NOVX expressing construct are useful to

SQ Sequence 5691 BP; 1318 A; 1658 C; 1574 G; 1141 T; 0 U; 0 Other;

Query Match 100.0%; Score 5691; DB 6; Length 5691;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 5691; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAGCCATGACCTGAGACCTGCTTCTTCCACCTCTTCATGTTGGGATG 60
 DB 1 ATGAAGCCATGACCTGAGACCTGACCTTCTTCCACCTCTTCATGTTGGGATG 60
 QY 61 GGGCTCCGACCTTGTCTACCCGAGAGCAGCCCGCTGTGCCAGAGCAGCGTCAATT 120
 DB 61 GGGCTCCGACCTTGTCTACCCGAGAGCAGCCCGCTGTGCCAGAGCAGCGTCAATT 120

QY 121 GTCAATTCGAGAGAGAGCCCGGAGGTTTCAATCACTGGTGGATGAGAGACA 180
DB 121 GTCAATTCGAGAGAGAGCCCGGAGGTTTCAATCACTGGTGGATGAGAGACA 180
QY 181 GGAACATTTACTTGGGGGCGGTCAATCGGATTTACAGAGCTCCAGAGACTGAAGTC 240
DB 181 GGAACATTTACTTGGGGGCGGTCAATCGGATTTACAGAGCTCCAGAGACTGAAGTC 240
QY 241 TTGGTGCGCATGAGACAGAGGCGGAGAGAGACACCCCAAGTTTAAACCCCGGACATC 300
DB 241 TTGGTGCGCATGAGACAGAGGCGGAGAGAGACACCCCAAGTTTAAACCCCGGACATC 300
QY 301 GTCCAGACCTGCAATGAGACCTGACACCAACAAATGTCAACAAAGTCTCTCATAT 360
DB 301 GTCCAGACCTGCAATGAGACCTGACACCAACAAATGTCAACAAAGTCTCTCATAT 360
QY 361 GACTACAGAGAGAAACAGGCTGATTTGCTGGAGGCTGTACCAAGGCACTGAGCTG 420
DB 361 GACTACAGAGAGAAACAGGCTGATTTGCTGGAGGCTGTACCAAGGCACTGAGCTG 420
QY 421 CTGAGGCTGAGAGACCTTCAAGCTGGGGAGGCTTATCATAGAGAGACCTATCTG 480
DB 421 CTGAGGCTGAGAGACCTTCAAGCTGGGGAGGCTTATCATAGAGAGACCTATCTG 480
QY 481 TCAGGTGTCAAAGAGCGGCTCAGTCTTTGAGGTGATCGTCTTACAGCAACTGGAT 540
DB 481 TCAGGTGTCAAAGAGCGGCTCAGTCTTTGAGGTGATCGTCTTACAGCAACTGGAT 540
QY 541 GACAAGCTGTTCAATGACACGAGTGGATGGAGAGCCGAGATTTTCCACCAATCTCC 600
DB 541 GACAAGCTGTTCAATGACACGAGTGGATGGAGAGCCGAGATTTTCCACCAATCTCC 600
QY 601 AGCCGAGAACTGACCAAGAACTCTGAGGCGGATGGCAATGTTCCGATACGCTTCCATGAT 660
DB 601 AGCCGAGAACTGACCAAGAACTCTGAGGCGGATGGCAATGTTCCGATACGCTTCCATGAT 660
QY 661 GAGTTCGTGGCTGATGATTAAGATCCCTTCCGACACTTCAACCAATCCTTGACTT 720
DB 661 GAGTTCGTGGCTGATGATTAAGATCCCTTCCGACACTTCAACCAATCCTTGACTT 720
QY 721 GATATCTACTATGTCTATGTTTGAAGTGGCACTTGTCTATCTTTTGAACCTTCAA 780
DB 721 GATATCTACTATGTCTATGTTTGAAGTGGCACTTGTCTATCTTTTGAACCTTCAA 780
QY 781 CCTGAGATGATGTCTCCACAGGCTCCACCAACAGAGAGAGGTGTATATCCAAAGCTC 840
DB 781 CCTGAGATGATGTCTCCACAGGCTCCACCAACAGAGAGAGGTGTATATCCAAAGCTC 840
QY 841 GTGAGGCTTTGCAAGAGAGACAGAGCTTCAACTCTATGTAAGAGGTCCTATGGCTGT 900
DB 841 GTGAGGCTTTGCAAGAGAGACAGAGCTTCAACTCTATGTAAGAGGTCCTATGGCTGT 900
QY 901 GAGCGCACTGGAGTGAAGTACCGCTGTGAGGCTGCTTACTGTCCAAAGCGGGGCTC 960
DB 901 GAGCGCACTGGAGTGAAGTACCGCTGTGAGGCTGCTTACTGTCCAAAGCGGGGCTC 960
QY 961 GTGCTTGGCAGAGACCTTGGAGTCCATCCAGATATAGACCTGCTTTCACCGCTTCTCC 1020
DB 961 GTGCTTGGCAGAGACCTTGGAGTCCATCCAGATATAGACCTGCTTTCACCGCTTCTCC 1020
QY 1021 AAGGCGCAGAGACCGGAAATCCCTGATAGTCCGCTGTGATCTTCACTTCTTG 1080
DB 1021 AAGGCGCAGAGACCGGAAATCCCTGATAGTCCGCTGTGATCTTCACTTCTTG 1080
QY 1081 AAGCAGATTAATGACCGGATTAAGAGAGGCTGAGTCTTTTACCGGGGCGAGGAGAG 1140
DB 1081 AAGCAGATTAATGACCGGATTAAGAGAGGCTGAGTCTTTTACCGGGGCGAGGAGAG 1140
QY 1141 CTGAGCTGGCTGGCTCAAGGTGAGAGACATCCCTGACAGCACTGGCTTAAACCAT 1200
DB 1141 CTGAGCTGGCTGGCTCAAGGTGAGAGACATCCCTGACAGCACTGGCTTAAACCAT 1200
QY 1201 GACGATTAACCTTGTGGCTGAGACATGAATGCTCCCTGGGAGTGTCCGACATGTGTGCT 1260

DB 1201 GACGATTAACCTTGTGGCTGAGACATGAATGCTCCCTGGGAGTGTCCGACATGTGTGCT 1260
QY 1261 GGAATTCCTGCTTCAAGAGAGACAGGACCGGATGAGAGTGTGATCATCGATATGTCTAC 1320
DB 1261 GGAATTCCTGCTTCAAGAGAGACAGGACCGGATGAGAGTGTGATCATCGATATGTCTAC 1320
QY 1321 AAGAACAACCTCTGAGCTTGTGGGACCAAAAGTGGCAAGCTTGAAGAAATCCGGGTG 1380
DB 1321 AAGAACAACCTCTGAGCTTGTGGGACCAAAAGTGGCAAGCTTGAAGAAATCCGGGTG 1380
QY 1381 GATGAGCCCAAGGAGCAAGCCCTCAGTATGAGACGTGTGAGAGTGTGAGACCCCGGCCCA 1440
DB 1381 GATGAGCCCAAGGAGCAAGCCCTCAGTATGAGACGTGTGAGAGTGTGAGACCCCGGCCCA 1440
QY 1441 GTCCCTCCGGATATGAGCTTCTCCAAGAGCACAGAGCAACTCTACATATGTCAAGAGAG 1500
DB 1441 GTCCCTCCGGATATGAGCTTCTCCAAGAGCACAGAGCAACTCTACATATGTCAAGAGAG 1500
QY 1501 CAGCTCACAGAGTCCCTGTGAGTCTGTGATGATATCAGAGCTGCGGAGTCTCTT 1560
DB 1501 CAGCTCACAGAGTCCCTGTGAGTCTGTGATGATATCAGAGCTGCGGAGTCTCTT 1560
QY 1561 GGGCTCAGGCGAACCCCACTGTGGGTGTGTGCTGCAACAAGTGACACCGGAAAGAG 1620
DB 1561 GGGCTCAGGCGAACCCCACTGTGGGTGTGTGCTGCAACAAGTGACACCGGAAAGAG 1620
QY 1621 CGGTGTGACGGTCCAGAGAGACCCCGAGTTTCTCTGGAGATGAACAGATGTGTCCG 1680
DB 1621 CGGTGTGACGGTCCAGAGAGACCCCGAGTTTCTCTGGAGATGAACAGATGTGTCCG 1680
QY 1681 CTGACGGTTCATCCCAAAATATCTCCGCTCTCAGTATCAAGCGTGTGCTGTCTGAG 1740
DB 1681 CTGACGGTTCATCCCAAAATATCTCCGCTCTCAGTATCAAGCGTGTGCTGTCTGAG 1740
QY 1741 ACGTCAATATGCCGAGACTGTGAGCTGTGAGCTGTGAGCACTTGTGAGAGCTGTCAAG 1800
DB 1741 ACGTCAATATGCCGAGACTGTGAGCTGTGAGCTGTGAGCACTTGTGAGAGCTGTCAAG 1800
QY 1801 ATGATGAGGCTGTGTGTGGGCAATCAGATCCAGTGTCTCTCCCTGACCCAGAGAGTG 1860
DB 1801 ATGATGAGGCTGTGTGTGGGCAATCAGATCCAGTGTCTCTCCCTGACCCAGAGAGTG 1860
QY 1861 CCCGGAATCATCAGAGATGAGGAGCCACCATGTGTGATCAGCTTCAAAATCAAG 1920
DB 1861 CCCGGAATCATCAGAGATGAGGAGCCACCATGTGTGATCAGCTTCAAAATCAAG 1920
QY 1921 GAGACCGGAGATGACTTGTGCAAGCACAGCTTGTCTTCAATGTGAGAGCTGTCAAAAT 1980
DB 1921 GAGACCGGAGATGACTTGTGCAAGCACAGCTTGTCTTCAATGTGAGAGCTGTCAAAAT 1980
QY 1981 TGTGCTGTGTCTGTGTGAGAGTCCATACCGCTGCACTGTGATTAATCCGGATGTC 2040
DB 1981 TGTGCTGTGTCTGTGTGAGAGTCCATACCGCTGCACTGTGATTAATCCGGATGTC 2040
QY 2041 TGCACCCATGACCCCAAGACCTGCTCTTCAAGAAAGCCCAAGTGTGAGTGTGAGATC 2100
DB 2041 TGCACCCATGACCCCAAGACCTGCTCTTCAAGAAAGCCCAAGTGTGAGTGTGAGATC 2100
QY 2101 TGGCCCCAGCTGTGAGTGAAGATCCGTGGGCGCGTGAAGGTATCAAGGCTATC 2160
DB 2101 TGGCCCCAGCTGTGAGTGAAGATCCGTGGGCGCGTGAAGGTATCAAGGCTATC 2160
QY 2161 ACGTGAAGGCCAAGAACTTCCCGAGCCCAAGTGTGAGGCTGTGAGATGATC 2220
DB 2161 ACGTGAAGGCCAAGAACTTCCCGAGCCCAAGTGTGAGGCTGTGAGATGATC 2220
QY 2221 CTCAACATTCAGGAGAGAGAGAGAGTGCCTGCGCTTCAACAGCTCCAGAGTA 2280
DB 2221 CTCAACATTCAGGAGAGAGAGAGAGTGCCTGCGCTTCAACAGCTCCAGAGTA 2280
QY 2281 CAGTGCAGAACACCTCTTATTTCTATGAAGGATGAGATCAACACCTGCGGTGAG 2340

Db 2281 CAGTCCAGAACACTCTTATTCTATGAGGATGAGATCAACACTGCCCCGAG 2340
Qy 2341 TTGACAGTGTGTGAAATGGGCACTTCAACATTGACAAACCGCTCAGATTAAGTTAC 2400
Db 2341 TTGACAGTGTGTGAAATGGGCACTTCAACATTGACAAACCGCTCAGATTAAGTTAC 2400
Qy 2401 CTCTCAAGTGTGAGGCAATGGGTGAGAGCTGCGGCTGTGCTCAAGCTGACCAAG 2460
Db 2401 CTCTCAAGTGTGAGGCAATGGGTGAGAGCTGCGGCTGTGCTCAAGCTGACCAAG 2460
Qy 2461 TTGCAATGTGTGAGGCAATGGGTGAGAGCTGCGGCTGTGCTCAAGCTGACCAAG 2520
Db 2461 TTGCAATGTGTGAGGCAATGGGTGAGAGCTGCGGCTGTGCTCAAGCTGACCAAG 2520
Qy 2521 CAGGAGAGCCAGTGTGAGAGCTGTGCTGCAAAAGCAAGTGTGCAAAACCCCGCATC 2580
Db 2521 CAGGAGAGCCAGTGTGAGAGCTGTGCTGCAAAAGCAAGTGTGCAAAACCCCGCATC 2580
Qy 2581 ACAGAGATATCCCGGTGACAGGCCCCCGGAAAGGGGGACCAAGGTCACTATCCGAGG 2640
Db 2581 ACAGAGATATCCCGGTGACAGGCCCCCGGAAAGGGGGACCAAGGTCACTATCCGAGG 2640
Qy 2641 GAGAACCTGGGCTGTGAAATTTGCGGACATGCGCTCCCATGTCAAGGTGTGGCTGAG 2700
Db 2641 GAGAACCTGGGCTGTGAAATTTGCGGACATGCGCTCCCATGTCAAGGTGTGGCTGAG 2700
Qy 2701 TGCAGCCCTTATGTGATGATTACATCCCTGACAAACAGATGCTGTGAGATGGGGAG 2760
Db 2701 TGCAGCCCTTATGTGATGATTACATCCCTGACAAACAGATGCTGTGAGATGGGGAG 2760
Qy 2761 GCCAAGCCCAAGCAGATGAGAGCTGTGAGAGATTTGCTGCTGTGTGTGCTGAA 2820
Db 2761 GCCAAGCCCAAGCAGATGAGAGCTGTGAGAGATTTGCTGCTGTGTGTGCTGAA 2820
Qy 2821 TTGATGGCCCGGCTTCAACAGCTCTATTAATTGATGACACTGACTCTCAGATCTGAG 2880
Db 2821 TTGATGGCCCGGCTTCAACAGCTCTATTAATTGATGACACTGACTCTCAGATCTGAG 2880
Qy 2881 CCCAGCCGGGGCCCAATGTCCGAGGAGACCAAGTACATCAAGGCACTGAT 2940
Db 2881 CCCAGCCGGGGCCCAATGTCCGAGGAGACCAAGTACATCAAGGCACTGAT 2940
Qy 2941 GCGGAGAGCAAGTGTGTGATGTTGAAAGCAGCCGTCTTCAAGGAGAT 3000
Db 2941 GCGGAGAGCAAGTGTGTGATGTTGAAAGCAGCCGTCTTCAAGGAGAT 3000
Qy 2941 GCGGAGAGCAAGTGTGTGATGTTGAAAGCAGCCGTCTTCAAGGAGAT 3000
Db 2941 GCGGAGAGCAAGTGTGTGATGTTGAAAGCAGCCGTCTTCAAGGAGAT 3000
Qy 3001 CCAATCTTCAATTTGTCACACACACATCTCAGATGAGGTGTGAGATGAAGGTGCG 3060
Db 3001 CCAATCTTCAATTTGTCACACACACATCTCAGATGAGGTGTGAGATGAAGGTGCG 3060
Qy 3061 GTGAGGTGTGACAGGCGCAAGATCAACAGAGACCTGTCTTGAATGTGAAGCCCC 3120
Db 3061 GTGAGGTGTGACAGGCGCAAGATCAACAGAGACCTGTCTTGAATGTGAAGCCCC 3120
Qy 3121 ACCATCGTGTGAGATTTGAGCGAGATGAGATTTGTCAAGTGAACACACCCATCGCGTA 3180
Db 3121 ACCATCGTGTGAGATTTGAGCGAGATGAGATTTGTCAAGTGAACACACCCATCGCGTA 3180
Qy 3181 TGGGGAGCCCACTGTGACCTTCAACAGAACCCCAATCGGTGTCCAAAGCATGAGGAGAG 3240
Db 3181 TGGGGAGCCCACTGTGACCTTCAACAGAACCCCAATCGGTGTCCAAAGCATGAGGAGAG 3240
Qy 3241 GAGCAATCAATATCTGTGAGGTCTTGAAGCTTCAAGATGAGCTGTCAAGGCGCCGCG 3300
Db 3241 GAGCAATCAATATCTGTGAGGTCTTGAAGCTTCAAGATGAGCTGTCAAGGCGCCGCG 3300
Qy 3301 CTGCTTGTGGTCTTGAACACAGTCAAGCTTGAACGAGGCGCGAGAGATTGGCTTC 3360
Db 3301 CTGCTTGTGGTCTTGAACACAGTCAAGCTTGAACGAGGCGCGAGAGATTGGCTTC 3360
Qy 3361 ATCTGTGACAAAGTCAAGTCTGTCTCATCTCAACAGACCAACTTCACTACTATCC 3420
Db 3361 ATCTGTGACAAAGTCAAGTCTGTCTCATCTCAACAGACCAACTTCACTACTATCC 3420

Qy 3421 AACCCGTTGTTGAGGCTTTGGTCCCTCAGAGATCTGAGAGCTCAAGCTGACAGCC 3480
Db 3421 AACCCGTTGTTGAGGCTTTGGTCCCTCAGAGATCTGAGAGCTCAAGCTGACAGCC 3480
Qy 3481 ATCATCTTAAAGGCAAGAACTGATCCGCTGTGGCTGTGGGCGCAAGTGAAGCTGAAC 3540
Db 3481 ATCATCTTAAAGGCAAGAACTGATCCGCTGTGGCTGTGGGCGCAAGTGAAGCTGAAC 3540
Qy 3541 TACAGTGTGTGTTGGGAGAGAGCCGTGACCCGTGACCCGTGACATGATGCAAGCTGTC 3600
Db 3541 TACAGTGTGTGTTGGGAGAGAGCCGTGACCCGTGACCCGTGACATGATGCAAGCTGTC 3600
Qy 3601 TGGAGTCCCAACCTCATCGGACAGCAAAAGTATGAGCCGTGTGCTGAGTGAAG 3660
Db 3601 TGGAGTCCCAACCTCATCGGACAGCAAAAGTATGAGCCGTGTGCTGAGTGAAG 3660
Qy 3661 TACTCCCGGGATGTGTATTTGCCCCGAGACAGCCGCTGAGCCGCTGAGCTGTC 3720
Db 3661 TACTCCCGGGATGTGTATTTGCCCCGAGACAGCCGCTGAGCCGCTGAGCTGTC 3720
Qy 3721 AGCATGCAATGTGTGAGGCTGCTCTCATCATTTTCAATGCTGAGCCGTGCTAT 3780
Db 3721 AGCATGCAATGTGTGAGGCTGCTCTCATCATTTTCAATGCTGAGCCGTGCTAT 3780
Qy 3781 AAACGCAAGTCCCGCAAGTGAATGACCTGACGCTGAGAGCGGTGAGATGAGAGCAAC 3840
Db 3781 AAACGCAAGTCCCGCAAGTGAATGACCTGACGCTGAGAGCGGTGAGATGAGAGCAAC 3840
Qy 3841 CTGAGTCCCGGTGAGGCTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3900
Db 3841 CTGAGTCCCGGTGAGGCTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3900
Qy 3901 CATGAGCTGACAGAGTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3960
Db 3901 CATGAGCTGACAGAGTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3960
Qy 3961 AACATGAGGAGTGTGCTTCCAGAGATTTGAAGACCACTGTCTCCGAGACTTGAAGT 4020
Db 3961 AACATGAGGAGTGTGCTTCCAGAGATTTGAAGACCACTGTCTCCGAGACTTGAAGT 4020
Qy 4021 CCGGAGTACCGGAG 4080
Db 4021 CCGGAGTACCGGAG 4080
Qy 4081 AACAGAGTGTGCTGCTTCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4140
Db 4081 AACAGAGTGTGCTGCTTCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4140
Qy 4141 GCGAGCGTGTGACAGTGTGCTGCTCATCATGAGAGAGAGAGAGAGAGAGAGAGAG 4200
Db 4141 GCGAGCGTGTGACAGTGTGCTGCTCATCATGAGAGAGAGAGAGAGAGAGAGAGAG 4200
Qy 4201 GCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4260
Db 4201 GCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4260
Qy 4261 AACACCTTAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4320
Db 4261 AACACCTTAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4320
Qy 4321 TGGTTTACTTCTCTCTCAAAATTTCTCAAGAGAGTGTGAGAGAGAGAGAGAGAGAG 4380
Db 4321 TGGTTTACTTCTCTCTCAAAATTTCTCAAGAGAGTGTGAGAGAGAGAGAGAGAGAG 4380
Qy 4381 CTGTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4440
Db 4381 CTGTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4440
Qy 4441 GCGGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4500
Db 4441 GCGGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4500

QY	4501	GTCTGTGAGCTGTGTGACGCCAGACAATGGCCAAAGCCCCAGAGTCCCAAGTAAACAATCTCTC	4550
Db	4501	GTCTGTGAGCTGTGTGACGCCAGACAATGGCCAAAGCCCCAGAGTCCCAAGTAAACAATCTCTC	4550
QY	4561	AACTGTGACACCACTCACTCAGGTCAAGAGAAATTCTGATGCGCATCTTTCAGATATGTG	4620
Db	4561	AACTGTGACACCACTCACTCAGGTCAAGAGAAATTCTGATGCGCATCTTTCAGATATATGTG	4620
QY	4621	CCTTGCTCCCAACCGGGCCCAAGCTGCAGATATATGATCTCGAGTGGGCGACAAGAAAGTGGG	4680
Db	4621	CCTTGCTCCCAACCGGGCCCAAGCTGCAGATATGATCTCGAGTGGGCGACAAGAAAGTGGG	4680
QY	4681	GCAAGGATGATCTTTCAGAGATGAAGACATCAACCAACAAGATTGAAATGATTGGAAACGA	4740
Db	4681	GCAAGGATGATCTTTCAGAGATGAAGACATCAACCAACAAGATTGAAATGATTGGAAACGA	4740
QY	4741	CTGAACACACTGSGCCCACTTACCAGGTGCAGATGGTTCCTGCTGGTGGCATTTATGTCTCAAG	4800
Db	4741	CTGAACACACTGSGCCCACTTACCAGGTGCAGATGGTTCCTGCTGGTGGCATTTATGTCTCAAG	4800
QY	4801	CAGGTGACAGCTTATTAACGACAGTGAACAACCTCCACCGTCTGCAGACCTCAGCAAGTAAA	4860
Db	4801	CAGGTGACAGCTTATTAACGACAGTGAACAACCTCCACCGTCTGCAGACCTCAGCAAGTAAA	4860
QY	4861	TATGAAAACATGATCCGGTACACCGGCGAGCCCCGACAGCTTCGCTCAGGACCTATG	4920
Db	4861	TATGAAAACATGATCCGGTACACCGGCGAGCCCCGACAGCTTCGCTCAGGACACCTATG	4920
QY	4921	ATCAATCTCTGACCTGGAGAGATGGAGTCAAGATGTGGCACTTAAGTGAAGAACACAGAGAC	4980
Db	4921	ATCAATCTCTGACCTGGAGAGATGGAGTCAAGATGTGGCACTTAAGTGAAGAACACAGAGAC	4980
QY	4981	GGAGACCGAAGAGAGGGGGACCGGGGGAGCAAGATGTGTGAATTCATCTGACCCGA	5040
Db	4981	GGAGACCGAAGAGAGGGGGACCGGGGGAGCAAGATGTGTGTGAATTCATCTGACCCGA	5040
QY	5041	CTCTCGGCACTTAAGGACACTGAGAAATTGTGGATGACCTTTGACACCATCTCTC	5100
Db	5041	CTCTCGGCACTTAAGGAGGACACTGAGAAATTGTGGATGACCTTTTGAACAATCTTC	5100
QY	5101	AGCAGCGGACACCGGTGCTCTGCCCCCTGCGCCATCAAGTACATGTTTGACTTCTCG	5160
Db	5101	AGCAGCGGACACCGGTGCTCTGCCCCCTGCGCCATCAAGTACATGTTTGACTTCTCG	5160
QY	5161	GATGAGCAGGCTGATAAACATGGCATTTATGACCCGACGTCCTCCGCACTACTGGAAAGAC	5220
Db	5161	GATGAGCAGGCTGATAAACATGGCATTTATGACCCGACGTCCTCCGCACTACTGGAAAGAC	5220
QY	5221	AATGACCTGCCCCCTAGGTTTGGGTCAACATGATCAGAAGAACCCGCACTTTGTGTGTGAC	5280
Db	5221	AATGACCTGCCCCCTAGGTTTGGGTCAACATGATCAGAAGAACCCGCACTTTGTGTGTGAC	5280
QY	5281	ATCCTAAGAACAGCATCAACAGCGCTGCTCTGTGTGTGCTCAGACTTTCATGTGAC	5340
Db	5281	ATCCTAAGAACAGCATCAACAGCGCTGCTCTGTGTGTGCTCAGACTTTCATGTGAC	5340
QY	5341	TCTTGCTCAAGTCAAGACACCGGCTGGGGACAAGACTGCGCTCCAAACAGCTGCTGAT	5400
Db	5341	TCTTGCTCAAGTCAAGACACCGGCTGGGGACAAGACTGCGCTCTCAACAGCTGCTGAT	5400
QY	5401	GCCAAAGACATCCCGAGCTACAAAGATTGGGTGAAGGATATTACTCAACATAGGAAG	5460
Db	5401	GCCAAAGACATCCCGAGCTACAAAGATTGGGTGAAGGATATTACTCAACATAGGAAG	5460
QY	5461	ATGCCAGCATTAAGGACCAACACATGAACGATATCTGGCTGAGACAGATCCCGATGAC	5520
Db	5461	ATGCCAGCATTAAGGACCAACACATGAACGATATCTGGCTGAGACAGATCCCGATGAC	5520
QY	5521	ATGATGAGTTCAACACCATGATGACACTCTCAGAGATCTTCTCTATGTGGCAAAATAC	5580
Db	5521	ATGATGAGTTCAACACCATGATGACACTCTCAGAGATCTTCTCTATGTGGCAAAATAC	5580
QY	5581	AGCGAGAGATCTTGGACTCTTGCACCAAGTACAGATGTGGGAACGAGAAATTGGCC	5640

Db	5581	AGCGAGGAGATCCTTGGACCTCTGGACCAAGATGACCGAGTGTGGAGACAGAAACTG3CC	5640
Qy	5641	TACGAACCTAGACAGACATCAACCCCTCATAGGCTTAGACAGCTGAATTA	5691
Db	5641	TACCAACTAGAACAGATCATTAACCTCTCATGAGCTTAGACAGCTGAATTA	5691
RESULT 2			
ID	ABA00062		
AB	ABA00062	standard; cDNA; 6367 BP.	
XX	ABA00062;		
XX	25-OCT-2002	(first entry)	
XX	CADHP-9	coding sequence, Incyte ID No: 7156379CB1.	
XX	Gene;	human; cell adhesion protein; CADHP; AIDS; Alzheimer's disease;	
XX	acquired immunodeficiency syndrome; thymic dysplasia; epilepsy;		
XX	renal tubular acidosis; congenital glaucoma; cancer; atherosclerosis;		
XX	Parkinson's disease; ss.		
XX	Homo sapiens.		
XX	Key	Location/Qualifiers	
XX	CDS	566..6250	
XX		/*tag = a	
XX		/product= "CADHP-9"	
XX	MO200259312-A2.		
XX	01-AUG-2002.		
XX	18-DEC-2001;	2001WO-US049206.	
XX	18-DEC-2000;	2000US-0256542P.	
XX	22-DEC-2000;	2000US-0259504P.	
XX	05-JAN-2001;	2001US-0260101P.	
XX	(INCYTE -)	INCYTE GENOMICS INC.	
XX	Duggan BM, Xu Y, Lee EA, Lee S, Lu DM, Warren BA, Yue H;		
XX	Gitzen KJ, Honchell CD, Burford N, Baughn MR, Tang TY, Hillman JL;		
XX	Gandhi AR, Kallik DA, Bandman O, Graul RC, Walla NK, Lu Y;		
XX	Ramkumar J, Yao MG, Lal PG;		
XX	WPI; 2002-590826/63.		
XX	P-PSDB; AAG79420.		
XX	New human cell adhesion proteins (CADHP) useful for treating, diagnosing		
XX	and preventing diseases or conditions associated with the aberrant CADHP		
XX	expression e.g. cancer, acquired immunodeficiency syndrome, Alzheimer's		
XX	disease and epilepsy.		
XX	Claim 5; Page 146-48; 149pp; English.		
XX	The sequences given in ABA00054-63 encode novel human cell adhesion		
XX	proteins (CADHP). The CADHP polypeptides and polynucleotides are useful		
XX	in treating, diagnosing and preventing diseases or conditions associated		
XX	with the decreased expression or overexpression of CADHP e.g. immune		
XX	system (acquired immunodeficiency syndrome, thymic dysplasia,		
XX	neurological (Alzheimer's disease, Parkinson's disease, epilepsy),		
XX	developmental (renal tubular acidosis, congenital glaucoma) and cell		
XX	proliferative (cancer, atherosclerosis) disorders. They are also useful		
XX	in assessing the effects of exogenous compounds on the expression of		
XX	nucleic acid and amino acid sequences of CADHP. The CADHP or its		
XX	fragments are useful in screening compounds for effectiveness as agonist		
XX	or antagonist of the polypeptides, or in altering the expression of the		
XX	target polynucleotide and compounds that specifically bind to or modulate		
XX	the activity of the polypeptide. The protein encoded by this cDNA		
XX	sequence shows homology to mouse plexin-2		

QY 2101 TCCCCCAGCTCTGAGATGGAACAAGATCTGTGTCCTGAGGAGGTATCAAGCTTATC 2160
Db 2666 TCCCCCAGCTCTGAGATGGAACAAGATCTGTGTCCTGAGGAGGTATCAAGCTTATC 2725
QY 2161 AGCTGGAAGCCAGAACTCCCCCAGGCCGAGTGGGAGAGGCTAAGAAATGCATC 2220
Db 2726 AGCTGGAAGCCAGAACTCCCCCAGGCCGAGTGGGAGAGGCTAAGAAATGCATC 2785
QY 2221 CTCAACATTCAGGGCAGCAGCAGCAGAGTCCCGCTTGCTCAACAGCTCCAGCGTA 2280
Db 2786 CTCAACATTCAGGGCAGCAGCAGCAGAGTCCCGCTTGCTCAACAGCTCCAGCGTA 2845
QY 2281 CAGTGCAGAACCTCTTATTCTTATGAAAGGATGGAATCAACAACCTGCGCGTGGAG 2340
Db 2846 CAGTGCAGAACCTCTTATTCTTATGAAAGGATGGAATCAACAACCTGCGCGTGGAG 2905
QY 2341 TTGACAGTGTGATGGAATGGGCACTTCAACATTCACACCCAGCTCAGATTAAGTTGAC 2400
Db 2906 TTGACAGTGTGATGGAATGGGCACTTCAACATTCACACCCAGCTCAGATTAAGTTGAC 2965
QY 2401 CTCTACAGTGTGAGACCATGCTTGAGAGCTGCGGCTGTGCTCAAGGCTACCCAGAC 2460
Db 2966 CTCTACAGTGTGAGACCATGCTTGAGAGCTGCGGCTGTGCTCAAGGCTACCCAGAC 3025
QY 2461 TTGCGATGTGAGTGTGAGCCAGGGCCAGGCGAGTGCACCTGCGCGAGCACTGCGCTGCC 2520
Db 3026 TTGCGATGTGAGTGTGAGCCAGGGCCAGGCGAGTGCACCTGCGCGAGCACTGCGCTGCC 3085
QY 2521 CAGAGAGCCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2580
Db 3086 CAGAGAGCCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3145
QY 2581 ACAGAGATATCCCGGAGCAGGCCCCCGGAGAGGGGGCACCAGGCTCACTATCCAGAGG 2640
Db 3146 ACAGAGATATCCCGGAGCAGGCCCCCGGAGAGGGGGCACCAGGCTCACTATCCAGAGG 3205
QY 2641 GAGAACCTGAGGCTGGAATTCGCGACATGCTCCCATGTCAAGTTCTGCGCTGAGAG 2700
Db 3206 GAGAACCTGAGGCTGGAATTCGCGACATGCTCCCATGTCAAGTTCTGCGCTGAGAG 3265
QY 2701 TGCAAGCTTTAATGATGATGATTAATCCCTGAGACAGATGATGATGATGATGATGATGATG 2760
Db 3266 TGCAAGCTTTAATGATGATGATTAATCCCTGAGACAGATGATGATGATGATGATGATGATG 3325
QY 2761 GCGAAGCCAGCAGCAGTGCAGGCTTCGTGAGATCTGCTGAGCTGTGTGTGCGCTGAA 2820
Db 3326 GCGAAGCCAGCAGCAGTGCAGGCTTCGTGAGATCTGCTGAGCTGTGTGTGCGCTGAA 3385
QY 2821 TTCAATGCGCGGTCTCTCAAGCTCTATTACTTCAATGACACTGACTCTCTCAATGATCAGAG 2880
Db 3386 TTCAATGCGCGGTCTCTCAAGCTCTATTACTTCAATGACACTGACTCTCTCAATGATCAGAG 3445
QY 2881 CCCAGCGGGGGCCCATGTCGCGAGGGAGCCCAAGTGCATCAAGGCAACCACTGAAT 2940
Db 3446 CCCAGCGGGGGCCCATGTCGCGAGGGAGCCCAAGTGCATCAAGGCAACCACTGAAT 3505
QY 2941 GCGGGAAGCAAGTGTGTGTGTATTTTGGAAAGAGGCTGTCTCTTCCCAAGGCAATCT 3000
Db 3506 GCGGGAAGCAAGTGTGTGTGTATTTTGGAAAGAGGCTGTCTCTTCCCAAGGCAATCT 3565
QY 3001 CCATCCATCATTTCTGTGAACACACACATCTCAGATAGGTGTAGATGAAGTGTGTG 3060
Db 3566 CCATCCATCATTTCTGTGAACACACACATCTCAGATAGGTGTAGATGAAGTGTGTG 3625
QY 3061 GTGAGAGTGAACAGGCGCAAGATCAACAGAGACTGTCTTCAATGTGTGAAGAGCC 3120
Db 3626 GTGAGAGTGAACAGGCGCAAGATCAACAGAGACTGTCTTCAATGTGTGAAGAGCC 3685
QY 3121 ACCATCGTGGGATTTGAGCCAGATGAGCATGTGTGAGGAACAACCATGCGCGTA 3180
Db 3686 ACCATCGTGGGATTTGAGCCAGATGAGCATGTGTGAGGAACAACCATGCGCGTA 3745
QY 3181 TGGGGAGCCCACTGGAACCTCATACAGAACCCCAAGATCGTGTCCAAAGCATGAGGGAG 3240

Db 3746 TGGGGAGCCCACTGGAACCTCATACAGAACCCCAAGATCGTGTCCAAAGCATGAGGGAG 3805
QY 3241 GAGACATCAATATCTGTAGGTTCTTGAAAGCTACTAGATGACCTGTGACAGCGCCGCG 3300
Db 3806 GAGACATCAATATCTGTAGGTTCTTGAAAGCTACTAGATGACCTGTGACAGCGCCGCG 3865
QY 3301 CTGCTGTGTGTCTTGAACCAACAGTCAAGCTTGAACCGAGAGGCCGAGAGATTGAGCTTC 3360
Db 3866 CTGCTGTGTGTCTTGAACCAACAGTCAAGCTTGAACCGAGAGGCCGAGAGATTGAGCTTC 3925
QY 3361 ATCTGGAACAAGTCCAGTCCCTGCTCATCTCAACAGAACCACTTCACTTATCC 3420
Db 3926 ATCTGGAACAAGTCCAGTCCCTGCTCATCTCAACAGAACCACTTCACTTATCC 3985
QY 3421 AACCGGTGTTTGAAGCTTTGAGTCCCTGAGAAATCCTGAGCTCAAGCTGCGACGCGC 3480
Db 3986 AACCGGTGTTTGAAGCTTTGAGTCCCTGAGAAATCCTGAGCTCAAGCTTGAAGCTGCGAC 4045
QY 3481 ATCATCTTAAAGGCAAGAACTGATCCGCTGTGCTGAGGCGCAACGTGAAGCTGAAC 3540
Db 4046 ATCATCTTAAAGGCAAGAACTGATCCGCTGTGCTGAGGCGCAACGTGAAGCTGAAC 4105
QY 3541 TACACTGTGTGTGTTGGGAGAGAGCCGTGACCGTGAACCGTTCAGATGTTCAGCTGCTC 3600
Db 4106 TACACTGTGTGTGTTGGGAGAGAGCCGTGACCGTGAACCGTTCAGATGTTCAGCTGCTC 4165
QY 3601 TGGCAGTCCCCCAACTCATGCGAGGCAACAAAGTGAATGACCGCTGTGCTGATGAGAG 3660
Db 4166 TGGCAGTCCCCCAACTCATGCGAGGCAACAAAGTGAATGACCGCTGTGCTGATGAGAG 4225
QY 3661 TACTCCCCGGGAGTGTGTATCTTCCCGGACAGCCCGCTCAAGCTGCCCGCATGCTC 3720
Db 4226 TACTCCCCGGGAGTGTGTATCTTCCCGGACAGCCCGCTCAAGCTGCCCGCATGCTC 4285
QY 3721 AGCATGAGTGTGAGTGTGAGGAGGCTCTCATCATTTTCACTGTGAGCTCATTTGCTAT 3780
Db 4286 AGCATGAGTGTGAGTGTGAGGAGGCTCTCATCATTTTCACTGTGAGCTCATTTGCTAT 4345
QY 3781 AAAGCGAAGTCCCGGAAAGTGAACCTCAACGCTGAAGCGGCTGAGATGAGATGAGAAC 3840
Db 4346 AAAGCGAAGTCCCGGAAAGTGAACCTCAACGCTGAAGCGGCTGAGATGAGATGAGAAC 4405
QY 3841 CTGAGTCCGAGTGTGAGGCTCTGAGTGAAGAGACCTTGTGCGAGCTGAGACGGAATC 3900
Db 4406 CTGAGTCCGAGTGTGAGGCTCTGAGTGAAGAGACCTTGTGCGAGCTGAGACGGAATC 4465
QY 3901 CATGAGTGAACAGTGAACCTGATGAGAGCCGGGATTCGTTCTTGAGCTATAGAACTTAC 3960
Db 4466 CATGAGTGAACAGTGAACCTGATGAGAGCCGGGATTCGTTCTTGAGCTATAGAACTTAC 4525
QY 3961 ACCATGCGGAGTGTGTTTCCCGAGAAATGAGAACCAACCTGTCTCCGAGACCTTGAAGTC 4020
Db 4526 ACCATGCGGAGTGTGTTTCCCGAGAAATGAGAACCAACCTGTCTCTCCGAGACCTTGAAGTC 4585
QY 4021 CCGGCTTACCGGAGAGAGCTGTGAGAAAGGCTGAAGCTTTGTGCCAGCTCATCAAC 4080
Db 4586 CCGGCTTACCGGAGAGAGCTGTGAGAAAGGCTGAAGCTTTGTGCCAGCTCATCAAC 4645
QY 4081 AACAGGTGTGTCTGTCTCTTCAATCCGAGAGCTTGAAGCTCCAGGAGCTTCCCATG 4140
Db 4646 AACAGGTGTGTCTGTCTCTTCAATCCGAGAGCTTGAAGCTCCAGGAGCTTCCCATG 4705
QY 4141 CGCGACGTTGGAAGTGTGAGCTCACTCATATGACCGTGTGAGAGCAAGCTGAGTAC 4200
Db 4706 CGCGACGTTGGAAGTGTGAGCTCACTCATATGACCGTGTGAGAGCAAGCTGAGTAC 4765
QY 4201 GCGACTGATGTGTAAGCAGTGTGAGCGGACCTCATTTGACAAAGACCTGAGAGCAAG 4260
Db 4766 GCGACTGATGTGTAAGCAGTGTGAGCGGACCTCATTTGACAAAGACCTGAGAGCAAG 4825
QY 4261 AACCAACCTTAAAGCTGTCTCAGAGAGCTGAGTCAATGATGATGATGATGATGATGATGATG 4320

DR P-PSDB; ABG34075.
 XX New PRO polypeptides and polynucleotides encoding the polypeptides.
 PT useful in gene therapy, chromosome identification, tissue typing, or for
 PS genetic analysis of individuals with genetic disorders.
 XX

Claim 2; Fig 91; 218bp; English.

CC This invention relates to the cDNA and protein sequences of novel
 CC secreted and transmembrane polypeptides PRO polypeptides. The invention
 CC also comprises a method for producing the proteins of the invention by
 CC recombinant means and antibodies specific for the protein of the
 CC invention. The antibody may be used for detecting the PRO proteins of the
 CC invention and may be used to modify their activity. polynucleotides may
 CC be used as hybridisation probes for a cDNA library to isolate the full-
 CC length PRO cDNA or to isolate other cDNAs, to construct hybridisation
 CC probes for mapping the gene which encodes that PRO and for genetic
 CC analysis of individuals with genetic disorders, in assays to identify
 CC other proteins or molecules involved in binding reaction, to generate
 CC transgenic animals or knock-out animals which in turn are useful in the
 CC development and screening of therapeutically useful reagents, for
 CC chromosome identification, and tissue typing. The PRO polypeptides are
 CC useful in gene therapy, and as molecular weight markers for protein
 CC electrophoresis purposes. The sequences may also be used to detect
 CC overexpression on PRO polypeptides in cancerous tumours and for screening
 CC for differentially expressed genes using microarray technology. The
 CC present sequence represents a cDNA encoding a human PRO protein of the
 CC invention
 XX

XX Sequence 2597 BP; 688 A; 570 C; 642 G; 697 T; 0 U; 0 Other;

XX Query Match 13.9%; Score 793; DB 6; Length 2597;

XX Best Local Similarity 100.0%; Pred. No. 0;

XX Matches 793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 579 CGAGTATTTTCCACATCTCCAGCGGAAATGACCAAGAACTCTGAGGCGGATGGCAT 638
 DB 1 CGAGTATTTTCCACATCTCCAGCGGAAATGACCAAGAACTCTGAGGCGGATGGCAT 60
 QY 639 GTTCGCGTACGCTCTTCATGATGAGTTCGTGCGCTGATGATTAATCCCTTCGACAC 698
 DB 61 GTTCGCGTACGCTCTTCATGATGAGTTCGTGCGCTGATGATTAATCCCTTCGACAC 120
 QY 699 CTTACCATCATCCCTGACTTTGATCTACTATGTCATGCTTATGAGAGTGGCAACTT 758
 DB 121 CTTACCATCATCCCTGACTTTGATCTACTATGTCATGCTTATGAGAGTGGCAACTT 180
 QY 759 TGTCTACTTTTGAACCTCCAACTGATGATGTCCTCCACAGGCTCCACCAAGA 818
 DB 181 TGTCTACTTTTGAACCTCCAACTGATGATGTCCTCCACAGGCTCCACCAAGA 240
 QY 819 GCAGGTGATATCAATCAAGCTCGTGAAGCTTTGCAAGAGACAGACCTTCAACTCTA 878
 DB 241 GCAGGTGATATCAATCAAGCTCGTGAAGCTTTGCAAGAGACAGACCTTCAACTCTA 300
 QY 879 TGTAGAGGTGCGCATTTGCTGTGAGCGCAGTGGGCTGAGTACCGCTGCTGAGGCTGC 938
 DB 301 TGTAGAGGTGCGCATTTGCTGTGAGCGCAGTGGGCTGAGTACCGCTGCTGAGGCTGC 360
 QY 939 CTACCTGTCCAAAGCGGGGCGCTGCTGAGAGACCTTGGATCCATCAAGATGATGA 998
 DB 361 CTACCTGTCCAAAGCGGGGCGCTGCTGAGAGACCTTGGATCCATCAAGATGATGA 420
 QY 999 CTTCTCTCTTCAACCGCTTCTTCCAAAGGCGCAGAAACCGGAAATGATCCCTGATGATGC 1058
 DB 421 CTTCTCTCTTCAACCGCTTCTTCCAAAGGCGCAGAAACCGGAAATGATCCCTGATGATGC 480
 QY 1059 GAGCGCTGTGCATCTTCAATCTTGAAGCAGATTAATGACCGCTTGAAGAGGCGCTGACATC 1118
 DB 481 GAGCGCTGTGCATCTTCAATCTTGAAGCAGATTAATGACCGCTTGAAGAGGCGCTGACATC 540
 QY 1119 TTGTATCCGGGCGAGGCGACGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 1178

DB 541 TTGTATCCGGGCGAGGCGACGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 600
 QY 1179 CAGCAGTGCCTCTTAAACATGACATTAATCTTGTGCGCTGAGACATGAATGCTCCCT 1238
 DB 601 CAGCAGTGCCTCTTAAACATGACATTAATCTTGTGCGCTGAGACATGAATGCTCCCT 660
 QY 1239 GCGAGTGTCCGACATGCTGTGGAATCCCGCTTCAAGAGAGAGAGGCGCATGAC 1298
 DB 661 GCGAGTGTCCGACATGCTGTGGAATCCCGCTTCAAGAGAGAGAGGCGCATGAC 720
 QY 1299 GTCTGTATGCGATATGCTTACAGAAACCACTCTGCTGCTTGTGGGACCAAAAGTG 1358
 DB 721 GTCTGTATGCGATATGCTTACAGAAACCACTCTGCTGCTTGTGGGACCAAAAGTG 780
 QY 1359 CAAGCTGAAGAG 1371
 DB 781 CAAGCTGAAGAG 793
 RESULT 4
 ADA01359
 ID ADA01359 standard; cDNA; 2597 BP.
 XX ADA01359;
 AC
 AC
 DT 06-NOV-2003 (first entry)
 XX
 D3 Human PRO polynucleotide #46.
 XX
 KW Human; gene; ss; PRO; secreted polypeptide; transmembrane polypeptide;
 KW tumour necrosis factor-alpha; TNF-alpha; blood; chondrocyte cell; tumour;
 KW adrenal; lung; colon; breast; prostate; rectum; cervix; liver; cancer;
 KW microvascular endothelial cell; endothelial cell tube formation;
 KW sports-related joint problem; articular cartilage defect; osteoarthritis;
 KW rheumatoid arthritis; osteopathic; anti-rheumatic; antiarthritic.
 XX
 OS Homo sapiens.
 XX
 PN US2003068779-A1.
 XX
 PD 10-APR-2003.
 XX
 PF 16-SEP-2002; 2002US-00245107.
 XX
 PR 09-MAY-2001; 2001US-0290589P.
 PR 29-AUG-2001; 2001WO-US027099.
 PR 18-JUL-2002; 2002US-00197942.
 XX
 PA (GENTH) GENTECH INC.
 XX
 PI Baker KP, Bacon DL, Filvaroff E, Goddard A, Grimaldi JC,
 PI Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WT, Zhang Z,
 PI Fong S;
 PI
 XX
 DR MPI: 2003-625484/59.
 DR P-PSDB; ADA01360.
 XX
 XX Novel isolated PRO1313, PRO20080 or PRO21383 polypeptide useful for
 PT stimulating proliferation of human microvascular endothelial cells, and
 PT PRO6018 polypeptide useful for stimulating proliferation of chondrocyte
 PT cells.
 PS
 PS Claim 2; Fig 91; 307bp; English.
 CC The invention relates to isolated human PRO polypeptides (secreted and
 CC transmembrane polypeptides) and the polynucleotides encoding them. The
 CC invention also relates to an antibody which specifically binds to a PRO
 CC polypeptide, a method for stimulating the release of tumour necrosis
 CC factor-alpha (TNF-alpha) from human blood, a method for stimulating the
 CC proliferation or differentiation of chondrocyte cells and a method for
 CC detecting the presence of a tumour in a mammal (e.g. adrenal, lung,
 CC colon, breast, prostate, rectal, cervical and liver tumours). The
 CC polynucleotides are useful in molecular biology, including uses as

hybridisation probes, in chromosome and gene mapping, in generating antisense RNA and DNA and in gene therapy. The polynucleotides may also be used in preparing PRO polypeptides by recombinant techniques and in generating either transgenic animals or knock-out animals which are useful in the development and screening of therapeutically useful reagents. The PRO polypeptides or antibodies are used in preparing a medicament for treating a condition responsive to the polypeptides or antibodies, such as tumours, for stimulating and inhibiting proliferation of human microvascular endothelial cells, for inducing endothelial cell tube formation and for treating sports-related joint problems, articular cartilage defects, osteoarthritis and rheumatoid arthritis. This sequence represents a human PRO polynucleotide of the invention.

Sequence 2597 BP; 688 A; 570 C; 642 G; 697 T; 0 U; 0 Other;

Query Match 13.9%; Score 793; DB 8; Length 2597;

Best Local Similarity 100.0%; Pred. No. 0; Matches 793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

579 CGAGTATTTTCCACCATCTCCAGCCGGAACCTGACCAAGACCTGAGGCGGATGGCAT 638
1 CGAGTATTTTCCACCATCTCCAGCCGGAACCTGACCAAGACCTGAGGCGGATGGCAT 60
639 GTTCCGCTACGCTCTCCATGATGAGTTGCTGAGCTCGATGATTAAGATCCCTTGGACAC 698
61 GTTCCGCTACGCTCTCCATGATGAGTTGCTGAGCTCGATGATTAAGATCCCTTGGACAC 120
699 GTTCCACCATCATCCTCTGACTTGTATATCTACTATGTCATAGTCTTTTACAGTGGCACTT 758
121 GTTCCACCATCATCCTCTGACTTGTATATCTACTATGTCATAGTCTTTTACAGTGGCACTT 180
759 TGTCTACTTTTGGACCCCTCCAACTGAGATGATGCTCTCCACAGGCTCCACCAAGAGA 818
181 TGTCTACTTTTGGACCCCTCCAACTGAGATGATGCTCTCCACAGGCTCCACCAAGAGA 240
819 GCAGGTGTATATCATCOAGCTCTGAGGCTTTGCAAGAGACACAGCTTCAACTCTCTA 878
241 GCAGGTGTATATCATCOAGCTCTGAGGCTTTGCAAGAGACACAGCTTCAACTCTCTA 300
879 TGTATAGAGGTGCCATTGGCTGTGAGCGAGTGGGGTGGAGTACCGCTGCTCAGGCTGC 938
301 TGTATAGAGGTGCCATTGGCTGTGAGCGAGTGGGGTGGAGTACCGCTGCTCAGGCTGC 360
939 CTACCTGTCTCAAAAGCGGGGCGGTGCTGGAGACCCCTTGGATTCATCCAGATGATGA 998
361 CTACCTGTCTCAAAAGCGGGGCGGTGCTGGAGACCCCTTGGATTCATCCAGATGATGA 420
999 CTGCTCTTCAACCGCTTTCTCCAAAGGCGCAAGGCGGAAATGAATCCCTGATGATGC 1058
421 CTGCTCTTCAACCGCTTTCTCCAAAGGCGCAAGGCGGAAATGAATCCCTGATGATGC 480
1059 GGGCCGTGTCATCTTCAATCTTGAAGCAGATAAATGACCGCATTAAGAGCGGCTGCACTC 1118
481 GGGCCGTGTCATCTTCAATCTTGAAGCAGATAAATGACCGCATTAAGAGCGGCTGCACTC 540
1119 TTGTTTACCGGGGAGAGGACCGTGGACCTGGCGCTGCAAGGAGTGAAGACATCCCGTG 1178
541 TTGTTTACCGGGGAGAGGACCGTGGACCTGGCGCTGCAAGGAGTGAAGACATCCCGTG 600
1179 CAGCAGTGGCGCTCTTAACCATTAAGCATTAAGTCTGTGAGCTGATGATGATGCTCCCT 1238
601 CAGCAGTGGCGCTCTTAACCATTAAGCATTAAGTCTGTGAGCTGATGATGATGCTCCCT 660
1239 GGGAGTGTCCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1298
661 GGGAGTGTCCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
1299 GTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1358
721 GTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
1359 CAACTGGAAGAG 1371

Db 781 CAACTGGAAGAG 793

RESULT 5
ADA43788
ADA43788 standard; cDNA; 2597 BP.
ADA43788;
20-NOV-2003 (first entry)

Human cDNA encoding secreted/transmembrane polypeptide PRO34003.
65; gene; human; PRO; secreted protein; transmembrane protein;
endothelial cell tube formation; chondrocyte cell differentiation;
microvascular endothelial cell; tumour; lung tumour; colon tumour;
breast tumour; prostate tumour; rectal tumour; kidney tumour;
liver tumour; cytostatic; vaccine.
Homo sapiens.
US2003064474-A1.
03-APR-2003.
16-SRP-2002; 2002US-00245859.
29-AUG-2001; 2001WO-US0247099.
18-JUL-2002; 2002US-00197942.
(GENTH) GENENTECH INC.
Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC,
Gurney AL, Smith V, Stephan JP, Watanabe CK, Wood WI, Zhang Z,
Fong S;
WPI; 2003-605867/57.
P-PSDB; ADA43789.
New isolated nucleic acid encoding a PRO polypeptide, e.g. PRO20080 or
PRO21383, useful in molecular biology, chromosome and gene mapping, in
generating antisense RNA and DNA, and in gene therapy.
Claim 2; Fig 91; 308pp; English.
The invention relates to an isolated secreted/transmembrane (PRO)
polypeptide, having at least 80% sequence identity to a sequence selected
from any one of the 57 amino acid sequences given in specification, or to
a sequence encoded by a nucleic acid molecule selected from any one of
the nucleic acids deposited under any of the ATCC accession numbers given
in specification, or a sequence having at least 80% identity to PRO
lacking its associated signal peptide, an extracellular domain of PRO
with or without its associated signal peptide. Also included are vectors,
transformed host cells, anti-PRO antibodies, the nucleic acids encoding
PRO, PRO fusion proteins, inducing endothelial cell tube formation (by
administering PRO281, PRO1560, PRO189, PRO4499, PRO6308, PRO6000,
PRO10275, PRO21207, PRO20933 or PRO34274 polypeptide or its agonist) and
an oligonucleotide probe derived from any one of the above nucleotide
sequences. PRO6018 polypeptide is useful for stimulating the
proliferation or differentiation of chondrocyte cells. PRO1313, PRO20080
and PRO21383 polypeptides are useful for stimulating the proliferation of
human microvascular endothelial cells. PRO6071, PRO4487 and PRO6006
polypeptides are useful for inhibiting the proliferation of human
microvascular endothelial cells. PRO polypeptides are useful for
detecting the presence of tumour in a mammal, including tumours of lung,
colon, breast, prostate, rectal, kidney and liver. PRO281, PRO1560,
PRO189, PRO4499, PRO6308, PRO6000, PRO10275, PRO21207, PRO20933 and
PRO34274 polypeptides are useful for inducing endothelial cell tube
formation. PRO or the antibody are useful in the preparation of a
medicament for treating a condition responsive to PRO polypeptide. The
oligonucleotide probes are useful for isolating genomic and cDNA
nucleotide sequences, for measuring or detecting the expression of an
associated gene, and as antisense probes. PRO nucleic acid is useful as a

CC hybridisation probe, in chromosome and gene mapping, in the generation of
CC antisense RNA and DNA, and for the preparation PRO polypeptides. PRO and
CC PRO nucleic acid are useful as therapeutic agents, e.g. vaccines. The
CC present sequence encodes a PRO protein.

XX Sequence 2597 BP; 688 A; 570 C; 642 G; 697 T; 0 U; 0 Other;

Query Match 13.9%; Score 793; DB 8; Length 2597;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 579 CGAGTATTTTCCACCATCTCCAGCCGGAAACTCAACCAAGACTCTTAGGGCGGATGGCAT 638
DB 1 CGAGTATTTTCCACCATCTCCAGCCGGAAACTCAACCAAGACTCTTAGGGCGGATGGCAT 60
QY 639 GTTCGCGTACGCTCTCCATGATGAGTTCGTGGCCCTCCATGATTAAATCCCTTCGACAC 698
DB 61 GTTCGCGTACGCTCTCCATGATGAGTTCGTGGCCCTCCATGATTAAATCCCTTCGACAC 120
QY 699 CTTCACCATCAATCCCTGACTTGAATCTACTATGCTATGCTATGCTATGCTATGCTATGCT 758
DB 121 CTTCACCATCAATCCCTGACTTGAATCTACTATGCTATGCTATGCTATGCTATGCTATGCT 180
QY 759 TGTCTACTTTTGAACCTCCACCTGAGATGATGCTGCTCCACAGGCTCCACCAAGAGA 818
DB 181 TGTCTACTTTTGAACCTCCACCTGAGATGATGCTGCTCCACAGGCTCCACCAAGAGA 240
QY 819 GCAGGTGATATCATCCCAAGCTCGTAGAGCTTTGCAAGAGACACAGCTTCAACTCTCTA 878
DB 241 GCAGGTGATATCATCCCAAGCTCGTAGAGCTTTGCAAGAGACACAGCTTCAACTCTCTA 300
QY 879 TGTAGAGTGGCCCATTTGCTGTGAGCGAGTGGAGTGAACCGCTGCGTGCAGGCTGC 938
DB 301 TGTAGAGTGGCCCATTTGCTGTGAGCGAGTGGAGTGAACCGCTGCGTGCAGGCTGC 360
QY 939 CTACCTGTCCAAAGCGGGGGCGGTGCTTGAGCAGNACCTTGGATCCATCCAGATGATGA 998
DB 361 CTACCTGTCCAAAGCGGGGGCGGTGCTTGAGCAGNACCTTGGATCCATCCAGATGATGA 420
QY 999 CCTGCTTTTACCGCTTCTTCCAGGGGCAAGCGGAAATGAAATCCCTGATGATGTC 1058
DB 421 CCTGCTTTTACCGCTTCTTCCAGGGGCAAGCGGAAATGAAATCCCTGATGATGTC 480
QY 1059 GGCCCTGTGCAATCTTCAATCTTGAAGCAGATAAATGACCGCATTAAGAGCGGCTGCAATC 1118
DB 481 GGCCCTGTGCAATCTTCAATCTTGAAGCAGATAAATGACCGCATTAAGAGCGGCTGCAATC 540
QY 1119 TTGTTACCGGGGGCGAGGGGACGCTGACCTGGCTTCAAGGTGAAGGACATCCCTG 1178
DB 541 TTGTTACCGGGGGCGAGGGGACGCTGACCTGGCTTCAAGGTGAAGGACATCCCTG 600
QY 1179 CAGCAGTGGGCTTAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1238
DB 601 CAGCAGTGGGCTTAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 1239 GGGAGTGTCCGACATGCTGCTGGAATTCCTTTCACGAGAGACAGGACCGCATGAC 1298
DB 661 GGGAGTGTCCGACATGCTGCTGGAATTCCTTTCACGAGAGACAGGACCGCATGAC 720
QY 1299 GTCTGTGATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1358
DB 721 GTCTGTGATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 1359 CAAAGTGAAGAG 1371
DB 781 CAAAGTGAAGAG 793

```

RESULT 6
ADA43556
ID ADA43556 standard; cDNA; 2597 BP.
XX AC ADA43556;

```

XX 20-NOV-2003 (first entry)
DT Human cDNA encoding secreted/transmembrane polypeptide PRO34003.
XX
XX
DE ss; gene; human; PRO; secreted protein; transmembrane protein;
XX microvascular endothelial cell; tumour; lung tumour; colon tumour;
XX liver tumour; prostate tumour; rectal tumour; kidney tumour;
XX Homo sapiens.
OS
XX US2003073196-A1.
XX
XX 17-APR-2003.
XX
XX 18-SEP-2002; 2002US-00246210.
XX
XX 04-APR-2001; 2001US-0282199P.
XX 29-ANG-2001; 2001MO-US027099.
XX 18-JUL-2002; 2002US-00197942.
XX
XX (GENT) GENENTECH INC.
XX Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC,
XX Gurney AL, Smith V, Stephan JP, Watanabe CK, Wood WI, Zhang Z;
XX Fong S;
XX WPI; 2003-743814/70.
XX P-PDB; ADA43557.
XX
XX New secreted and transmembrane PRO polypeptides, e.g. PRO20080 or
XX PRO21383 useful for stimulating the proliferation or differentiation of
XX chondrocyte cells and detecting the presence of a tumor in a mammal.
XX
XX Claim 2; Fig 91; 307pp; English.
XX
XX The invention relates to an isolated secreted/transmembrane (PRO)
XX polypeptide, having at least 80% sequence identity to a sequence selected
XX from any one of the 57 amino acid sequences given in specification, or to
XX a sequence encoded by a nucleic acid molecule selected from any one of
XX the nucleic acids deposited under any of the ATCC accession numbers given
XX in specification, or a sequence having at least 80% identity to PRO
XX lacking its associated signal peptide. Also included are vectors,
XX with or without its associated signal peptide. Also included are vectors,
XX transformed host cells, anti-PRO antibodies, the nucleic acids encoding
XX PRO, PRO fusion proteins, inducing endothelial cell tube formation (by
XX administering PRO281, PRO1560, PRO189, PRO4499, PRO6308, PRO6006,
XX PRO10275, PRO21207, PRO20933 or PRO34274 polypeptide or its agonist) and
XX an oligonucleotide probe derived from any one of the above nucleotide
XX sequences. PRO6018 polypeptide is useful for stimulating the
XX proliferation or differentiation of chondrocyte cells. PRO1313, PRO20080
XX and PRO21383 polypeptides are useful for stimulating the proliferation of
XX human microvascular endothelial cells. PRO6071, PRO4487 and PRO6006
XX polypeptides are useful for inhibiting the proliferation of human
XX microvascular endothelial cells. PRO polypeptides are useful for
XX detecting the presence of tumour in a mammal, including tumours of lung,
XX colon, breast, prostate, rectal, kidney and liver. PRO281, PRO1560,
XX PRO189, PRO4499, PRO6308, PRO6000, PRO10275, PRO21207, PRO20933 and
XX PRO34274 polypeptides are useful for inducing endothelial cell tube
XX formation. PRO or the antibody are useful in the preparation of a
XX medicament for treating a condition responsive to PRO polypeptide. The
XX oligonucleotide probes are useful for isolating genomic and cDNA
XX associated sequences, for measuring or detecting the expression of an
XX associated gene, and as antisense probes. PRO nucleic acid is useful as a
XX hybridisation probe, in chromosome and gene mapping, in the generation of
XX antisense RNA and DNA, and for the preparation PRO polypeptides. PRO and
XX PRO nucleic acid are useful as therapeutic agents, e.g. vaccines. The
XX present sequence encodes a PRO protein.
XX
XX Sequence 2597 BP; 688 A; 570 C; 642 G; 697 T; 0 U; 0 Other;

```

Query Match 13.9%; Score 793; DB 8; Length 2597;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 579 CGAGTATTTTCCACCATCTCCAGCCGGAACCTGACCAAGAACTCTGAGCGGATGGCAT 638
DB 1 CGAGTATTTTCCACCATCTCCAGCCGGAACCTGACCAAGAACTCTGAGCGGATGGCAT 60
QY 639 GTTGGCGTACGCTCTTCAATGATGATGGGCTCGATGATTAAGATCCCTTGGACAC 698
DB 61 GTTGGCGTACGCTCTTCAATGATGATGGGCTCGATGATTAAGATCCCTTGGACAC 120
QY 699 CTTCACCATCATCCCTGACTTGTATATCTACTATGTCTATGCTTTAGCACTGGCACTT 758
DB 121 CTTCACCATCATCCCTGACTTGTATATCTACTATGTCTATGCTTTAGCACTGGCACTT 180
QY 759 TGTCTACTTTTGAACCTCCACCTGAGATGGTGTCTCACAGGCTGCACCAAGGA 818
DB 181 TGTCTACTTTTGAACCTCCACCTGAGATGGTGTCTCACAGGCTGCACCAAGGA 240
QY 819 GAGGTGTATACATCCAGCTGTGAGGCTTTGCAAGAGACACAGCTTCAACTCTTA 878
DB 241 GAGGTGTATACATCCAGCTGTGAGGCTTTGCAAGAGACACAGCTTCAACTCTTA 300
QY 879 TGTAGAGTGGCCATTTGGCTGTGAGCGGATGGGCTGAGTAAACGCTCGACAGGCTGC 938
DB 301 TGTAGAGTGGCCATTTGGCTGTGAGCGGATGGGCTGAGTAAACGCTCGACAGGCTGC 360
QY 939 CTACCTGTCCAAAGCGGGGCGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 998
DB 361 CTACCTGTCCAAAGCGGGGCGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 999 CCTGCTCTTCAACCGCTTTCTCCAGGCGGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1058
DB 421 CCTGCTCTTCAACCGCTTTCTCCAGGCGGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 1059 GGGCCGTGTGCATCTTCAATGAGAGATGACCGCATTAAGAGAGAGAGAGAGAGAG 1118
DB 481 GGGCCGTGTGCATCTTCAATGAGAGATGACCGCATTAAGAGAGAGAGAGAGAGAG 540
QY 1119 TTGTACCGGGGCGAGGCGACGCTGAGACCTGGCTGCAAGGTGAAGAGAGAGAGAG 1178
DB 541 TTGTACCGGGGCGAGGCGACGCTGAGACCTGGCTGCAAGGTGAAGAGAGAGAGAG 600
QY 1179 CAGCAATGGGCTCTTAACCATTAAGATTAAGTGTGAGCATGATGCTCCCT 1238
DB 601 CAGCAATGGGCTCTTAACCATTAAGATTAAGTGTGAGCATGATGCTCCCT 660
QY 1239 GGGAGTGTCCGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1298
DB 661 GGGAGTGTCCGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
QY 1299 GTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1358
DB 721 GTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
QY 1359 CAAAGCTGAAGAG 1371
DB 781 CAAAGCTGAAGAG 793

```

RESULT 7

ADA01231 standard; cDNA; 2597 BP.

ADA01231;

06-NOV-2003 (first entry)

Human PRO polynucleotide #46.

Human; gene; ss; PRO; secreted polypeptide; transmembrane polypeptide;

tumour necrosis factor-alpha; TNF-alpha; blood; chondrocyte cell; tumour;

KW cancer; lung; colon; breast; prostate; rectum; kidney; liver;
 KM microvascular endothelial cell; endothelial cell tube formation.
 XX
 OS Homo sapiens.

US2003068782-A1.

10-APR-2003.

16-SEP-2002; 2002US-00245851.

27-APR-1999; 99US-0131271P.

22-OCT-1999; 99US-0162506P.

02-DEC-1999; 99WO-US028551.

29-AUG-2001; 2001WO-US027099.

18-JUL-2002; 2002US-00197942.

(GENTH) GENENTECH INC.

Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC,
 PI Garney AL, Smith V, Stephan JP, Watanabe CX, Wood WI, Zhang Z;
 PI Fong S;

WPI; 2003-625487/59.

P-PSDB; ADA01232.

Novel isolated PRO polypeptides e.g. PRO281 and PRO1560, useful in the
 PT preparation of a medicament for treating a condition responsive to PRO
 PT polypeptide, and as therapeutic agents e.g. vaccines.

Claim 2; Fig 91; 308bp; English.

XX The invention relates to isolated human PRO polypeptides (secreted and
 CC transmembrane polypeptides) and the polynucleotides encoding them. The
 CC invention also relates to an antibody which specifically binds to a PRO
 CC polypeptide, a method for stimulating the release of tumour necrosis
 CC factor-alpha (TNF-alpha) from human blood, a method for stimulating the
 CC proliferation or differentiation of chondrocyte cells and a method for
 CC detecting the presence of a tumour in a mammal (e.g. lung, colon, breast,
 CC prostate, rectal, kidney and liver tumours). The polynucleotides are
 CC useful in molecular biology, including uses as hybridisation probes, in
 CC chromosome and gene mapping, in generating antisense RNA and DNA and in
 CC gene therapy. The polynucleotides may also be used in preparing PRO
 CC polypeptides by recombinant techniques and in generating either
 CC transgenic animals or knock-out animals which are useful in the
 CC development and screening of therapeutically useful reagents. The PRO
 CC polypeptides or antibodies are used in preparing a medicament for
 CC treating a condition responsive to the polypeptides or antibodies, such
 CC as tumours, for stimulating and inhibiting proliferation of human
 CC microvascular endothelial cells and for inducing endothelial cell tube
 CC formation. This sequence represents a human PRO polynucleotide of the
 CC invention.

Sequence 2597 BP; 688 A; 570 C; 642 G; 697 T; 0 U; 0 Other;

Query Match 13.9%; Score 793; DB 8; Length 2597;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 579 CGAGTATTTTCCACCATCTCCAGCCGGAACCTGACCAAGAACTCTGAGCGGATGGCAT 638
DB 1 CGAGTATTTTCCACCATCTCCAGCCGGAACCTGACCAAGAACTCTGAGCGGATGGCAT 60
QY 639 GTTGGCGTACGCTCTTCAATGATGATGGGCTCGATGATTAAGATCCCTTGGACAC 698
DB 61 GTTGGCGTACGCTCTTCAATGATGATGGGCTCGATGATTAAGATCCCTTGGACAC 120
QY 699 CTTCACCATCATCCCTGACTTGTATATCTACTATGTCTATGCTTTAGCACTGGCACTT 758
DB 121 CTTCACCATCATCCCTGACTTGTATATCTACTATGTCTATGCTTTAGCACTGGCACTT 180
QY 759 TGTCTACTTTTGAACCTCCACCTGAGATGGTGTCTCACAGGCTGCACCAAGGA 818

```


QY 879 TGTAGAGGTGCCCATTTGGCTGTGAGCGAGTGGGGTGGAGTACCGCTGTCGACAGGCTGC 938
 DB 301 TGTAGAGGTGCCCATTTGGCTGTGAGCGAGTGGGGTGGAGTACCGCTGTCGACAGGCTGC 360
 QY 939 CTACCTGTCCAAAGGG 998
 DB 361 CTACCTGTCCAAAGGG 420
 QY 999 CCTGCTCTTCCACCGCTCTTCTCCAGAGGCGAGAGCGGAAATGAAATCCCTGATGATGC 1058
 DB 421 CCTGCTCTTCCACCGCTCTTCTCCAGAGGCGAGAGCGGAAATGAAATCCCTGATGATGC 480
 QY 1059 GGGCCTGTGCATCTTTCATTTTGAAGCAGATTAATATACCGCATTAAGAGCGGCTGCAATC 1118
 DB 481 GGGCCTGTGCATCTTTCATTTTGAAGCAGATTAATATACCGCATTAAGAGCGGCTGCAATC 540
 QY 1119 TTGTTACCGGGGGGAGAGGAGCGAGTGGCTGAGTGGCTGAGTGAAGAGCATCCCTG 1178
 DB 541 TTGTTACCGGGGGGAGAGGAGCGAGTGGCTGAGTGGCTGAGTGAAGAGCATCCCTG 600
 QY 1179 CAGCAGTGGCTCTTTCATTTTGAAGCAGATTAATATACCGCATTAAGAGCGGCTGCAATC 1238
 DB 601 CAGCAGTGGCTCTTTCATTTTGAAGCAGATTAATATACCGCATTAAGAGCGGCTGCAATC 660
 QY 1239 GGGAGTGGCTGAGTGGCTGAGTGGCTGAGTGGCTGAGTGGCTGAGTGGCTGAGTGGCTGAG 1298
 DB 661 GGGAGTGGCTGAGTGGCTGAGTGGCTGAGTGGCTGAGTGGCTGAGTGGCTGAGTGGCTGAG 720
 QY 1299 GTCTGTGATCGCATATGTTGTACAGAAACCACTCTCTGAGCTTTGTGGGACCAAAAGTGG 1358
 DB 721 GTCTGTGATCGCATATGTTGTACAGAAACCACTCTCTGAGCTTTGTGGGACCAAAAGTGG 780
 QY 1359 CAAGCTGAAGAG 1371
 DB 781 CAAGCTGAAGAG 793
 RESULT 9
 ADA43672 ID ADA43672 standard; cDNA; 2597 BP.
 AC ADA43672;
 DX 20-NOV-2003 (first entry)
 XX
 DE Human cDNA encoding secreted/transmembrane polypeptide PRO34003.
 XX
 KM ss: gene; human; PRO; secreted protein; transmembrane protein;
 KM endothelial cell tube formation; chondrocyte cell differentiation;
 KM microvascular endothelial cell; tumour; lung tumour; colon tumour;
 KM breast tumour; prostate tumour; rectal tumour; kidney tumour;
 KM liver tumour; cytosolic; vaccine.
 OS Homo sapiens.
 XX
 PN US2003073190-A1.
 XX
 PD 17-APR-2003.
 XX
 PF 09-SEP-2002; 2002US-00238283.
 XX
 PR 01-JUL-1998; 98US-0091358P.
 PR 02-JUN-1999; 99WO-US012252.
 PR 20-JUL-1999; 99US-0144758P.
 PR 28-JUL-1999; 99US-0146222P.
 PR 25-AUG-1999; 99US-00380137.
 PR 30-MAR-2000; 2000WO-US008439.
 PR 02-JUN-2000; 2000WO-US015264.
 PR 29-AUG-2001; 2001WO-US027099.
 PR 18-JUL-2002; 2002US-00197942.
 XX
 PA (GENTH) GENENTECH INC.
 XX

PI Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC,
 PI Gurney AL, Smith V, Stephan JP, Watanabe CK, Wood WI, Zhang Z,
 PI Fong S;
 XX
 DR WPI; 2003-585304/55.
 DR P-PSDB; ADA43673.
 PT New isolated nucleic acid encoding a PRO polypeptide, e.g. PRO20080 or
 PT PRO21383, useful in molecular biology, chromosome and gene mapping, in
 PT generating antisense RNA and DNA, and in gene therapy.
 XX
 PS Claim 2; Fig 91; 352pp; English.
 XX
 CC The invention relates to an isolated secreted/transmembrane (PRO)
 CC polypeptide, having at least 80% sequence identity to a sequence selected
 CC from any one of the 57 amino acid sequences given in specification, or to
 CC a sequence encoded by a nucleic acid molecule selected from any one of
 CC the nucleic acids deposited under any of the ATCC accession numbers given
 CC in specification, or a sequence having at least 80% identity to PRO
 CC lacking its associated signal peptide, an extracellular domain of PRO
 CC with or without its associated signal peptide. Also included are vectors,
 CC transformed host cells, anti-PRO antibodies, the nucleic acids encoding
 CC PRO, PRO fusion proteins, inducing endothelial cell tube formation (by
 CC administering PRO281, PRO1560, PRO189, PRO4499, PRO6308, PRO6000,
 CC PRO10275, PRO21207, PRO20993 or PRO34274 polypeptide or its agonist) and
 CC an oligonucleotide probe derived from any one of the above nucleotide
 CC sequences. PRO6018 polypeptide is useful for stimulating the
 CC proliferation or differentiation of chondrocyte cells. PRO1313, PRO20080
 CC and PRO21383 polypeptides are useful for stimulating the proliferation of
 CC human microvascular endothelial cells. PRO6071, PRO4487 and PRO6006
 CC polypeptides are useful for inhibiting the proliferation of human
 CC microvascular endothelial cells. PRO polypeptides are useful for
 CC detecting the presence of tumour in a mammal, including tumours of lung,
 CC colon, breast, prostate, rectal, kidney and liver. PRO281, PRO1560,
 CC PRO189, PRO4499, PRO6308, PRO6000, PRO10275, PRO20993 and
 CC PRO34274 polypeptides are useful for inducing endothelial cell tube
 CC formation. PRO or the antibody are useful in the preparation of a
 CC medicament for treating a condition responsive to PRO polypeptide. The
 CC oligonucleotide probes are useful for isolating genomic and cDNA
 CC nucleotide sequences, for measuring or detecting the expression of an
 CC associated gene, and as antisense probes. PRO nucleic acid is useful as a
 CC hybridisation probe, in chromosome and gene mapping, in the generation of
 CC antisense RNA and DNA, and for the preparation PRO polypeptides. PRO and
 CC PRO nucleic acid are useful as therapeutic agents, e.g. vaccines. The
 CC present sequence encodes a PRO protein.
 XX
 SQ Sequence 2597 BP; 688 A; 570 C; 642 G; 697 T; 0 U; 0 Other;
 Query Match 13.3%; Score 793; DB 8; Length 2597;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 579 CGAGTATTTTCCACCACTTCTCCAGCCGGAACCTGACCAAGAACTTGAGGCGATGGCAT 638
 DB 1 CGAGTATTTTCCACCACTTCTCCAGCCGGAACCTGACCAAGAACTTGAGGCGATGGCAT 60
 QY 639 GTTGGCGTACGTCCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 698
 DB 61 GTTGGCGTACGTCCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
 QY 699 CTTGACCATCATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 758
 DB 121 CTTGACCATCATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
 QY 759 TGTCTACTTTTGAACCTTCCAACTGAGTGTCTCCACCAAGGCTCCACCAAGGA 818
 DB 181 TGTCTACTTTTGAACCTTCCAACTGAGTGTCTCCACCAAGGCTCCACCAAGGA 240
 QY 819 GCAGGTGATATCATCAAGCTCGTAGGCTTTGCAAGAGAGACACAGCTTTCAACTCTTA 878
 DB 241 GCAGGTGATATCATCAAGCTCGTAGGCTTTGCAAGAGAGACACAGCTTTCAACTCTTA 300
 QY 879 TGTAGAGGTGCCCATTTGGCTGTGAGCGAGTGGGGTGGAGTACCGCTGTCGACAGGCTGC 938

QY	999	CCGCGCTTCAACCGCTTCTCCAAAGGCGAABAACGGAAAAATGAAATCCCTGGATGAGTC	1058
Db	421	CTGCTCTTCAACCGCTTCTCCAAAGGCGAABAATGAAATCCCTGGATGAGTC	480
QY	1059	GGCCTCTGTCATCTTCACTCTTGAAGCAGATAATATGACCGCATTTAAGAGCGCTGACGTC	1118
Db	481	GCGCCCTGTGCATCTTCACTCTTGAAGCAGATAAAGACCGCATTTAAGAGCGCTGACGTC	540
QY	1119	TTGTATCCGGGCGAGGCGACGCTGAGCCTGGCTGGCTCAAGGTGAAGACATCCCTG	1178
Db	541	TTGTATCCGGGCGAGGCGACGCTGAGCCTGGCTGGCTCAAGGTGAAGACATCCCTG	600
QY	1179	CAGCAGTGGGCTTTAACCATTGACAGATAACTTCTGTGGCTGGACATGAATGCTCCCT	1238
Db	601	CAGCAGTGGGCTTTAACCATTGACAGATAACTTCTGTGGCTGGACATGAATGCTCCCT	660
QY	1239	GGGAGTGTCCGACATGTCGCTGGAATTCGCCGCTTCAACGAGAGCAAGGACCGCATGAC	1298
Db	661	GGGAGTGTCCGACATGTCGCTGGAATTCGCCGCTTCAACGAGAGCAAGGACCGCATGAC	720
QY	1299	GTCTGTCATGCAATATGCTACAAACAACACTCTCTGAGCTTTGTGGGACCAAAAAGTGG	1358
Db	721	GTCTGTCATGCAATATGCTACAAACAACACTCTCTGAGCTTTGTGGGACCAAAAAGTGG	780
QY	1359	CAAGCTGAAGAAG 1371	
Db	781	CAAGCTGAAGAAG 793	

RESULT 11	
ADAO84422	
ID	ADAO84422 standard; cDNA; 2597 BP.
XX	
AC	ADAO84422;
XX	
DT	06-NOV-2003 (first entry)
XX	
DE	Novel human secreted and transmembrane protein PRO34003 cDNA.
XX	
XX	ss; gene; osteopathic; antirheumatic; antiarthritic; gene therapy;
KW	cell proliferation stimulator;
KW	chondrocyte cell differentiation stimulator;
KW	secreted and transmembrane protein; PRO; human; PRO1313; PRO20080;
KW	PRO2183; human microvascular endothelial cell proliferation; PRO6771;
KW	PRO4467; PRO6006; PRO240; PRO256; PRO698; PRO1002; PRO4316; tumour;
KW	adrenal tumour; lung tumour; colon tumour; breast tumour;
KW	prostate tumour; rectal tumour; cervical tumour; liver tumour;
KW	sports-related joint problem; articular cartilage defect; osteoarthritis;
KW	rheumatoid arthritis; tissue typing.
XX	
OS	Homo sapiens.
XX	
PN	US2003066783-A1.
XX	
PD	10-APR-2003.
XX	
PF	16-SEP-2002; 2002US-00245883.
XX	
ER	09-MAY-2001; 2001US-0290589P.
ER	23-AUG-2001; 2001WO-US027099.
PR	18-JUL-2002; 2002US-00197942.
XX	
PA	(GETH) GENENTECH INC.
XX	
PI	Baker KP, Baton DL, Filvaroff E, Goddard A, Grimaud JC;
PI	Gurney AL, Smith V, Stephan DT, Watanabe CK, Wood WI, Zhang Z;
PI	Pong S;
XX	
DR	WPI; 2003-625488/59.
DR	P-PSDB; ADAO84423.
XX	
XX	Novel isolated PRO1313, PRO20080 or PRO21383 polypeptide useful for
XX	stimulating proliferation of human microvascular endothelial cells, and

PT	PRO6018	polypeptide	useful	for	stimulating	proliferation	of	chondrocyte	PT
PR	cells.								PR
PS	Claim 2;	Fig 91;	308bp;	English.					PS
XX									XX
CC	The invention	describes	an	isolated	PRO	(secreted	and	transmembrane)	CC
CC	polypeptide	(I).	PRO6018	polypeptide	is	useful	for	stimulating	CC
CC	proliferation	or	differentiation	of	chondrocyte	cells.	PRO1333,	PRO20080,	CC
CC	or	PRO23383	polypeptide	is	useful	for	stimulating	the	CC
CC	proliferation	of	human	microvascular	endothelial	cells.	PRO6071,	PRO4487,	CC
CC	or	PRO6006	polypeptide	is	useful	for	inhibiting	the	CC
CC	proliferation	of	human	microvascular	endothelial	cells.	PRO6071,	PRO4487,	CC
CC	or	PRO6006	polypeptide	is	useful	for	inhibiting	the	CC
CC	proliferation	of	human	microvascular	endothelial	cells.	PRO6071,	PRO4487,	CC
CC	or	PRO6006	polypeptide	is	useful	for	inhibiting	the	CC
CC	proliferation	of	human	microvascular	endothelial	cells.	PRO6071,	PRO4487,	CC
CC	or	PRO6006	polypeptide	is	useful	for	inhibiting	the	CC
CC	proliferation	of	human	microvascular	endothelial	cells.	PRO6071,	PRO4487,	CC
CC	or	PRO6006	polypeptide	is	useful	for	inhibiting	the	CC
CC	proliferation	of	human	microvascular	endothelial	cells.	PRO6071,	PRO4487,	CC
CC	or	PRO6006	polypeptide	is	useful	for	inhibiting	the	CC
CC	proliferation	of	human	microvascular	endothelial	cells.	PRO6071,	PRO4487,	CC
CC	or	PRO6006	polypeptide	is	useful	for	inhibiting	the	CC
CC	proliferation	of	human	microvascular	endothelial	cells.	PRO6071,	PRO4487,	CC
CC	or	PRO6006	polypeptide	is	useful	for	inhibiting	the	CC
CC	proliferation	of	human	microvascular	endothelial	cells.	PRO6071,	PRO4487,	CC
CC	or	PRO6006	polypeptide	is	useful	for	inhibiting	the	CC
CC	proliferation	of	human	microvascular	endothelial	cells.	PRO6071,	PRO4487,	CC
CC	or	PRO6006	polypeptide	is	useful	for	inhibiting	the	CC
CC	proliferation	of	human	microvascular	endothelial	cells.	PRO6071,	PRO4487,	CC
CC	or	PRO6006	polypeptide	is	useful	for	inhibiting	the	CC
CC	proliferation	of	human	microvascular	endothelial	cells.	PRO6071,	PRO4487,	CC
CC	or	PRO6006	polypeptide	is	useful	for	inhibiting	the	CC
CC	proliferation	of	human	microvascular	endothelial	cells.	PRO6071,	PRO4487,	CC
CC	or	PRO6006	polypeptide	is	useful	for	inhibiting	the	CC
CC	proliferation	of	human	microvascular	endothelial	cells.	PRO6071,	PRO4487,	CC
CC	or	PRO6006	polypeptide	is	useful	for	inhibiting	the	CC
CC	proliferation	of	human	microvascular	endothelial	cells.	PRO6071,	PRO4487,	CC
CC	or	PRO6006	polypeptide	is	useful	for	inhibiting	the	CC
CC	proliferation	of	human	microvascular	endothelial	cells.	PRO6071,	PRO4487,	CC
CC	or	PRO6006	polypeptide	is	useful	for	inhibiting	the	CC
CC	proliferation	of	human	microvascular	endothelial	cells.	PRO6071,	PRO4487,	CC
CC	or	PRO6006	polypeptide	is	useful	for	inhibiting	the	CC
CC	proliferation	of	human	microvascular	endothelial	cells.	PRO6071,	PRO4487,	CC
CC	or	PRO6006	polypeptide	is	useful	for	inhibiting	the	CC
CC	proliferation	of	human	microvascular	endothelial	cells.	PRO6071,	PRO4487,	CC
CC	or	PRO6006	polypeptide	is	useful	for	inhibiting	the	CC
CC	proliferation	of	human	microvascular	endothelial	cells.	PRO6071,	PRO4487,	CC
CC	or	PRO6006	polypeptide	is	useful	for	inhibiting	the	CC
CC	proliferation	of	human	microvascular	endothelial	cells.	PRO6071,	PRO4487,	CC
CC	or	PRO6006	polypeptide	is	useful	for	inhibiting	the	CC
CC	proliferation	of	human	microvascular	endothelial	cells.	PRO6071,	PRO4487,	CC
CC	or	PRO6006	polypeptide	is	useful	for	inhibiting	the	CC
CC	proliferation	of	human	microvascular	endothelial	cells.	PRO6071,	PRO4487,	CC
CC	or	PRO6006	polypeptide	is	useful				

QY 933 CTACCTGTCCAAAGCGGGGCGCTGCTTGGCAGGACCCCTGGAGTCACATCCAGATGATGA 998
 Db 361 CTACCTGTCCAAAGCGGGGCGCTGCTTGGCAGGACCCCTGGAGTCACATCCAGATGATGA 420
 QY 999 CCTGCTCTTACCGGCTCTTCTCCAAAGGCGCAGAAAGCGAAATATCCCTGGATGATG 1058
 Db 421 CCTGCTCTTACCGGCTCTTCTCCAAAGGCGCAGAAATATCCCTGGATGATGATG 480
 QY 1059 GGGCCTGTGATCTTCACTTGAAGCAGATTAATGACCCCATTAAGGACCGGCTGAGTC 1118
 Db 481 GGGCCTGTGATCTTCACTTGAAGCAGATTAATGACCCCATTAAGGACCGGCTGAGTC 540
 QY 1119 TTGTACCGGGGCGAGGCGACGCTGACCTGCGCTGCTCAAGGTGAGAGACATCCCTG 1178
 Db 541 TTGTACCGGGGCGAGGCGACGCTGACCTGCGCTGCTCAAGGTGAGAGACATCCCTG 600
 QY 1179 CAGCAGTGGCTCTTAAACATTAACATTAATCTTGTGCTGACATGATGCTCCCT 1238
 Db 601 CAGCAGTGGCTCTTAAACATTAACATTAATCTTGTGCTGACATGATGCTCCCT 660
 QY 1239 GGGAGTGTCCGACATGCTGCTGAGTAATCCGCTTCCAGGAGACAGGACCGCATGAC 1298
 Db 661 GGGAGTGTCCGACATGCTGCTGAGTAATCCGCTTCCAGGAGACAGGACCGCATGAC 720
 QY 1299 GTCTGTATGCGCATATGCTTCAAGAAACCACTCTTGTGCTGCTGAGGACCAAAAGTGG 1358
 Db 721 GTCTGTATGCGCATATGCTTCAAGAAACCACTCTTGTGCTGCTGAGGACCAAAAGTGG 780
 QY 1359 CAAGCTGAAGAG 1371
 Db 781 CAAGCTGAAGAG 793

RESULT 12
 ID ADB9715 standard; cDNA; 2597 BP.
 XX ADB9715;
 AC ADB9715;
 XX 04-DEC-2003 (first entry)
 DT
 XX
 DE Human PRO polynucleotide SEQ ID 91.
 XX
 KM Human; gene; ss; PRO; secreted polypeptide; transmembrane polypeptide;
 KM tumour necrosis factor-alpha; TNF-alpha; blood; chondrocyte cell; tumour;
 KM adrenal; lung; colon; breast; prostate; rectum; cervix; liver; cancer;
 KM microvascular endothelial cell; endothelial cell tube formation;
 KM sports-related joint problem; articular cartilage defect; osteoarthritis;
 KM rheumatoid arthritis; osteopathic; antiinflammatory; antiarthritic.
 XX
 OS Homo sapiens.
 XX
 PN US2003082728-A1.
 XX
 PD 01-MAY-2003.
 XX
 PF 16-SEP-2002; 2002US-00245185.
 XX
 PR 07-JUL-1998; 98US-00919782.
 PR 02-JUN-1999; 99WO-US012252.
 PR 25-AUG-1999; 99US-00380137.
 PR 29-AUG-2001; 2001WO-US027099.
 PR 18-JUL-2002; 2002US-00197942.
 XX
 PA (GENENTECH INC.
 PI Baker KP, Baton DL, Filvaroff E, Goddard A, Grimaldi JC,
 PI Gurey AL, Smith V, Stephan JP, Watanabe CK, Wood WI, Zhang Z,
 PI Fong S,
 XX WPI: 2003-743698/70.
 DR P-Psdb; ADB9716.
 XX

PT New isolated, secreted and transmembrane PRO polypeptides and nucleic
 PT acids, useful for diagnosing, preventing and/or treating tumors, such as
 PT lung, colon, breast, prostate, rectal, kidney or liver tumors.
 XX
 XX Claim 2; Fig 91; 308pp; English.
 XX
 CC The invention relates to isolated human PRO polypeptides (secreted and
 CC transmembrane polypeptides) and the polynucleotides encoding them. The
 CC invention also relates to an antibody which specifically binds to a PRO
 CC polypeptide, a method for stimulating the release of tumour necrosis
 CC factor-alpha (TNF-alpha) from human blood, a method for stimulating the
 CC proliferation or differentiation of chondrocyte cells and a method for
 CC detecting the presence of a tumour in a mammal (e.g. adrenal, lung,
 CC colon, breast, prostate, rectal, cervical and liver tumors). The
 CC polynucleotides are useful in molecular biology, including uses as
 CC hybridisation probes, in chromosome and gene mapping, in generating
 CC antisense RNA and DNA and in gene therapy. The polynucleotides may also
 CC be used in preparing PRO polypeptides by recombinant techniques and in
 CC generating either transgenic animals or knock-out animals which are
 CC useful in the development and screening of therapeutically useful
 CC reagents. The PRO polypeptides or antibodies are used in preparing a
 CC medicament for treating a condition responsive to the polypeptides or
 CC antibodies, such as tumours, for stimulating and inhibiting proliferation
 CC of human microvascular endothelial cells, for inducing endothelial cell
 CC tube formation and for treating sports-related joint problems, articular
 CC cartilage defects, osteoarthritis and rheumatoid arthritis. This sequence
 CC represents a human PRO polynucleotide of the invention.
 CC
 SQ Sequence 2597 BP; 688 A; 570 C; 642 G; 697 T; 0 U; 0 Other;

Query Match 13.9%; Score 793; DB 8; Length 2597;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 579 CGAGTATTTTCCACCATCTCCAGCCGGAACCTGACCAAGACTGTGAGGGGATGGAT 638
 Db 1 CGAGTATTTTCCACCATCTCCAGCCGGAACCTGACCAAGACTGTGAGGGGATGGAT 60
 QY 639 GTTGGGAGTACGCTTCCATGATGATGCTGCTGAGGCTGATTAAGATCCCTTGAGCAC 698
 Db 61 GTTGGGAGTACGCTTCCATGATGATGCTGCTGAGGCTGATTAAGATCCCTTGAGCAC 120
 QY 699 CTTCACCATCATCCTGATCTTATATCTATATCTATGTTTACAGTGGCACTT 758
 Db 121 CTTCACCATCATCCTGATCTTATATCTATATCTATGTTTACAGTGGCACTT 180
 QY 759 TGTCTACTTTTGAACCTCCCAACCTGAGATGCTGCTCCACAGGCTCCACCAAGAGA 818
 Db 181 TGTCTACTTTTGAACCTCCCAACCTGAGATGCTGCTCCACAGGCTCCACCAAGAGA 240
 QY 819 GAGAGTATATACATCAAGCTGTGAGGCTTTCAGAGAGACACAGCCTTCACTCTTA 878
 Db 241 GAGAGTATATACATCAAGCTGTGAGGCTTTCAGAGAGACACAGCCTTCACTCTTA 300
 QY 879 TGTAGAGTGGCCATTTGGCTGAGACGAGTGGGGTGGAGTACCGCTGCTGAGCTGC 938
 Db 301 TGTAGAGTGGCCATTTGGCTGAGACGAGTGGGGTGGAGTACCGCTGCTGAGCTGC 360
 QY 939 CTACCTGTCCAAAGCGGGGCGCTGCTTGGCAGGACCCCTGGAGTCACATCCAGATGATGA 998
 Db 361 CTACCTGTCCAAAGCGGGGCGCTGCTTGGCAGGACCCCTGGAGTCACATCCAGATGATGA 420
 QY 999 CCTGCTCTTACCGGCTCTTCTCCAAAGGCGCAGAAAGCGAAATATCCCTGGATGATG 1058
 Db 421 CCTGCTCTTACCGGCTCTTCTCCAAAGGCGCAGAAAGCGAAATATCCCTGGATGATG 480
 QY 1059 GGGCCTGTGATCTTCACTTGAAGCAGATTAATGACCCCATTAAGGACCGGCTGAGTC 1118
 Db 481 GGGCCTGTGATCTTCACTTGAAGCAGATTAATGACCCCATTAAGGACCGGCTGAGTC 540
 QY 1119 TTGTACCGGGGCGAGGCGACGCTGACCTGCGCTGCTCAAGGTGAGAGACATCCCTG 1178
 Db 541 TTGTACCGGGGCGAGGCGACGCTGACCTGCGCTGCTCAAGGTGAGAGACATCCCTG 600

RESULT 14
AD866153 standard: cDNA; 2597 BP.
ID ADB66153
AC ADB66153;
XX
XX
DT 04-DEC-2003 (first entry)
XX
XX Human cDNA encoding secreted/transmembrane polypeptide PRO34003.
KW ss; gene; human; PRO; secreted protein; transmembrane protein;
KM endothelial cell tube formation; chondrocyte cell differentiation;
KM microvascular endothelial cell; tumour; lung tumour; colon tumour;
KM breast tumour; prostate tumour; rectal tumour; kidney tumour;
KM liver tumour; cytostatic; vaccine.
XX
XX Homo sapiens.
OS
XX
PN US2003082729-A1.
XX
PD 01-MAY-2003.
XX
PF 16-SEP-2002; 2002US-00245427.
XX
PR 02-FEB-2001; 2001US-0266421P.
PR 09-FEB-2001; 2001US-0267623P.
PR 29-AUG-2001; 2001WO-US027099.
PR 16-JUL-2002; 2002US-00197942.
XX
PA (GENTH) GENENTECH INC.
XX
XX Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC,
PI Gunney AL, Smith V, Stephan JP, Watanabe CK, Wood WT, Zhang Z,
PI Fong S,
XX
DR WPJ; 2003-786917/74.
DR P-PDSB; AD866154.
XX
XX New isolated, secreted and transmembrane PRO polypeptides and nucleic acids, useful for diagnosing, preventing and/or treating tumors, such as lung, colon, breast, prostate, rectal, kidney or liver tumors.
PS
PS Claim 2; Fig 91; 308bp; English.

The invention relates to an isolated secreted/transmembrane (PRO) polypeptide, having at least 80% sequence identity to a sequence selected from any one of the 57 amino acid sequences given in specification, or to a sequence encoded by a nucleic acid molecule selected from any one of the nucleic acids deposited under any of the ATCC accession numbers given in specification, or a sequence having at least 80% identity to PRO lacking its associated signal peptide, an extracellular domain of PRO with or without its associated signal peptide. Also included are vectors, transformed host cells, anti-PRO antibodies, the nucleic acids encoding PRO, PRO fusion proteins, inducing endothelial cell tube formation (by administering PRO261, PRO1560, PRO189, PRO4499, PRO6308, PRO6000, PRO10275, PRO21207, PRO20933 or PRO34274 polypeptide or its agonist) and an oligonucleotide probe derived from any one of the above nucleotide sequences. PRO6018 polypeptide is useful for stimulating the proliferation or differentiation of chondrocyte cells. PRO313, PRO30080 and PRO21383 polypeptides are useful for stimulating the proliferation of human microvascular endothelial cells. PRO6071, PRO4487 and PRO6006 polypeptides are useful for inhibiting the proliferation of human microvascular endothelial cells. PRO polypeptides are useful for detecting the presence of tumour in a mammal, including tumours of lung, colon, breast, prostate, rectal, kidney and liver. PRO261, PRO1560, PRO189, PRO4499, PRO6308, PRO6000, PRO10275, PRO21207, PRO20933 and PRO34274 polypeptides are useful for inducing endothelial cell tube formation. PRO or the antibody are useful in the preparation of a reagent for treating a condition responsive to PRO polypeptide. The oligonucleotide probes are useful for isolating genomic and cDNA nucleotide sequences, for measuring or detecting the expression of an associated gene, and as antisense probes. PRO nucleic acid is useful as a hybridisation probe, in chromosome and gene mapping, in the generation of

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2004, 02:26:30 ; Search time 260 Seconds
(without alignments)
12147.021 Million cell updates/sec

Title: US-09-964-956-12
Perfect score: 5691
Sequence: 1 atgaagacatgcctcgaa.....gtctagacatggaataaa 5691

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

Issued Parents NA: *
1: /cgn2_6/prodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/prodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/prodata/2/ina/PCFUS.COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	0.4	249	4	US-09-252-991A-3017
2	20	0.4	311	4	US-09-313-294A-7021
3	20	0.4	462	4	US-09-401-064-165
4	20	0.4	616	3	US-09-385-982-238
5	20	0.4	894	4	US-09-252-991A-2754
6	20	0.4	1029	4	US-09-252-991A-2651
7	20	0.4	1455	4	US-09-252-991A-3212
8	20	0.4	1456	4	US-09-576-594-779
9	20	0.4	1636	4	US-08-530-915A-31
10	20	0.4	2522	4	US-09-620-312D-702
11	20	0.4	13842	3	US-09-105-537-5
12	20	0.4	36778	3	US-09-105-537-5
13	20	0.4	38506	3	US-09-320-878-19
14	20	0.4	38506	4	US-09-141-908-1
15	20	0.4	38506	4	US-09-657-440-19
16	19	0.3	531	5	PCT-US95-04583-1
17	19	0.3	588	4	US-08-924-695A-3
18	19	0.3	791	4	US-09-657-346A-10
19	19	0.3	900	1	US-07-955-041-1
20	19	0.3	900	1	US-08-227-455-1
21	19	0.3	900	1	US-08-472-482-1
22	19	0.3	900	1	US-08-487-069-1
23	19	0.3	1236	4	US-09-328-352-3326
24	19	0.3	1407	4	US-09-894-844-35
25	19	0.3	2238	1	US-07-841-651-1
26	19	0.3	2477	4	US-09-907-794A-169
27	19	0.3	2477	4	US-09-905-125A-169

101	-17	0.3	401	4	US-09-339-338-196	Sequence 196, App	C 174	-17	0.3	148	4	US-09-485-632B-1	Sequence 1, Appl
102	-17	0.3	401	4	US-09-433-826B-196	Sequence 196, App	C 175	-17	0.3	153	4	US-09-472-112-2	Sequence 2, Appl
103	-17	0.3	401	4	US-09-604-287A-196	Sequence 196, App	C 176	-17	0.3	160	4	US-09-485-632B-7	Sequence 7, Appl
104	-17	0.3	401	4	US-09-834-759-196	Sequence 196, App	C 177	-17	0.3	177	4	US-09-252-991A-11683	Sequence 11683, A
105	-17	0.3	402	4	US-09-543-681A-3530	Sequence 3530, Ap	C 178	-17	0.3	1469	1	US-08-053-867A-3	Sequence 3, Appl
106	-17	0.3	417	4	US-09-621-976-1344	Sequence 1344, Ap	C 179	-17	0.3	1469	1	US-08-053-867A-6	Sequence 6, Appl
107	-17	0.3	423	4	US-09-252-991A-13079	Sequence 13029, A	C 180	-17	0.3	1512	4	US-09-540-236-1742	Sequence 1742, Ap
108	-17	0.3	453	4	US-09-621-976-1518	Sequence 1518, Ap	C 181	-17	0.3	1337	2	US-09-252-991A-11555	Sequence 11555, A
109	-17	0.3	474	4	US-09-621-976-17384	Sequence 17384, A	C 182	-17	0.3	1337	2	US-08-839-006-1	Sequence 1, Appl
110	-17	0.3	481	4	US-09-621-976-17383	Sequence 17383, A	C 183	-17	0.3	1339	2	US-09-089-584-55	Sequence 55, Appl
111	-17	0.3	484	4	US-09-621-976-2660	Sequence 2660, Ap	C 184	-17	0.3	1569	2	US-08-901-306-3	Sequence 3, Appl
112	-17	0.3	486	4	US-09-621-976-75	Sequence 76, Appl	C 185	-17	0.3	1569	2	US-09-180-271-3	Sequence 3, Appl
113	-17	0.3	497	4	US-09-621-976-736	Sequence 736, App	C 186	-17	0.3	1575	4	US-09-252-991A-1521	Sequence 1521, Ap
114	-17	0.3	500	4	US-09-833-381-1307	Sequence 1307, Ap	C 187	-17	0.3	1622	4	US-09-599-360B-43	Sequence 43, Appl
115	-17	0.3	505	3	US-08-808-303-11	Sequence 11, Appl	C 188	-17	0.3	1643	4	US-09-623-655-753	Sequence 753, Appl
116	-17	0.3	505	3	US-08-996-533-11	Sequence 11, Appl	C 189	-17	0.3	1656	4	US-09-489-039A-1878	Sequence 3878, Ap
117	-17	0.3	510	4	US-09-252-991A-15535	Sequence 15535, A	C 190	-17	0.3	1669	4	US-09-591-466C-1	Sequence 1, Appl
118	-17	0.3	515	4	US-09-621-976-3443	Sequence 3443, Ap	C 191	-17	0.3	1671	4	US-09-795-691-3	Sequence 3, Appl
119	-17	0.3	516	3	US-09-085-199B-16	Sequence 16, Appl	C 192	-17	0.3	1737	4	US-09-591-466C-3	Sequence 3, Appl
120	-17	0.3	520	2	US-08-691-814B-102	Sequence 102, App	C 193	-17	0.3	1747	1	US-08-176-620A-1	Sequence 1, Appl
121	-17	0.3	524	4	US-08-986-171E-323	Sequence 923, App	C 194	-17	0.3	1747	1	US-08-463-862-1	Sequence 1, Appl
122	-17	0.3	567	4	US-09-621-976-1059	Sequence 1059, Ap	C 195	-17	0.3	1747	2	US-08-461-985-1	Sequence 1, Appl
123	-17	0.3	573	4	US-09-205-258-84	Sequence 84, Appl	C 196	-17	0.3	1747	2	US-08-458-887-1	Sequence 1, Appl
124	-17	0.3	573	4	US-09-252-991A-3321	Sequence 3321, Ap	C 197	-17	0.3	1747	3	US-08-532-787B-1	Sequence 1, Appl
125	-17	0.3	577	4	US-08-801-742-2	Sequence 2, Appl	C 198	-17	0.3	1747	3	US-08-532-012C-1	Sequence 1, Appl
126	-17	0.3	585	4	US-09-340-620A-50	Sequence 50, Appl	C 199	-17	0.3	1872	4	US-09-280-116-110	Sequence 110, App
127	-17	0.3	588	1	US-08-808-303-1	Sequence 1, Appl	C 200	-17	0.3	1874	4	US-08-888-818C-1	Sequence 1, Appl
128	-17	0.3	588	1	US-08-808-303-2	Sequence 2, Appl	C 201	-17	0.3	1875	4	US-09-252-991A-5054	Sequence 5054, Ap
129	-17	0.3	588	3	US-08-996-533-1	Sequence 1, Appl	C 202	-17	0.3	1904	1	US-08-412-375A-5	Sequence 5, Appl
130	-17	0.3	588	3	US-08-996-533-2	Sequence 2, Appl	C 203	-17	0.3	1926	3	US-09-117-250-4	Sequence 4, Appl
131	-17	0.3	600	3	US-09-328-111-433	Sequence 433, App	C 204	-17	0.3	1962	4	US-09-352-991A-5520	Sequence 5520, Ap
132	-17	0.3	627	4	US-09-107-532A-1745	Sequence 1745, Ap	C 205	-17	0.3	1962	4	US-09-620-312D-715	Sequence 715, App
133	-17	0.3	656	3	US-08-998-416-56	Sequence 56, Appl	C 206	-17	0.3	1990	1	US-08-454-720D-35	Sequence 932, App
134	-17	0.3	694	4	US-09-976-594-359	Sequence 359, App	C 207	-17	0.3	1997	4	US-09-620-312D-932	Sequence 4, Appl
135	-17	0.3	697	4	US-09-328-111-686	Sequence 686, App	C 208	-17	0.3	1999	4	US-09-167-109-4	Sequence 12, Appl
136	-17	0.3	736	1	US-08-808-303-7	Sequence 7, Appl	C 209	-17	0.3	1999	4	US-09-921-099A-12	Sequence 1, Appl
137	-17	0.3	736	1	US-08-996-533-7	Sequence 7, Appl	C 210	-17	0.3	2003	4	US-08-691-814E-1	Sequence 231, App
138	-17	0.3	740	4	US-09-340-620A-48	Sequence 48, Appl	C 211	-17	0.3	2063	3	US-08-724-586-1	Sequence 1, Appl
139	-17	0.3	758	4	US-09-833-381-1308	Sequence 1308, Ap	C 212	-17	0.3	2066	3	US-09-421-632-1	Sequence 1, Appl
140	-17	0.3	774	4	US-09-352-991A-1601	Sequence 1601, Ap	C 213	-17	0.3	2086	4	US-09-932-190-1	Sequence 1, Appl
141	-17	0.3	774	4	US-09-352-991A-1601	Sequence 1601, Ap	C 214	-17	0.3	2086	4	US-09-932-190-1	Sequence 1, Appl
142	-17	0.3	779	4	US-09-623-655-5952	Sequence 595, App	C 215	-17	0.3	2091	1	US-08-655-640-1	Sequence 1, Appl
143	-17	0.3	780	4	US-09-452-991A-1652	Sequence 1652, Ap	C 216	-17	0.3	2117	2	US-08-655-640-3	Sequence 3, Appl
144	-17	0.3	780	4	US-09-489-039A-2658	Sequence 2658, Ap	C 217	-17	0.3	2145	4	US-09-366-257-1	Sequence 1, Appl
145	-17	0.3	791	4	US-08-936-165A-132	Sequence 132, App	C 218	-17	0.3	2165	4	US-09-220-132-61	Sequence 61, Appl
146	-17	0.3	804	4	US-09-489-039A-1865	Sequence 1865, Ap	C 219	-17	0.3	2177	4	US-09-016-434-1084	Sequence 1084, Ap
147	-17	0.3	828	4	US-09-352-991A-3347	Sequence 3347, Ap	C 220	-17	0.3	2229	4	US-09-252-991A-7718	Sequence 7718, Ap
148	-17	0.3	854	4	US-09-352-991A-5034	Sequence 5034, Ap	C 221	-17	0.3	2229	4	US-09-252-991A-5162	Sequence 5162, Ap
149	-17	0.3	948	4	US-09-489-039A-5573	Sequence 5573, Ap	C 222	-17	0.3	2319	1	US-08-154-019-1	Sequence 1, Appl
150	-17	0.3	1001	4	US-09-641-638-258	Sequence 258, App	C 223	-17	0.3	2319	1	US-08-461-333-1	Sequence 1, Appl
151	-17	0.3	1038	4	US-09-485-632B-3	Sequence 3, Appl	C 224	-17	0.3	2319	1	US-08-464-167-1	Sequence 1, Appl
152	-17	0.3	1039	4	US-09-665-189A-64	Sequence 64, Appl	C 225	-17	0.3	2319	3	US-09-158-313-1	Sequence 1, Appl
153	-17	0.3	1046	4	US-09-221-017B-496	Sequence 496, App	C 226	-17	0.3	2330	3	US-08-456-108-1	Sequence 1, Appl
154	-17	0.3	1096	4	US-09-716-161A-3	Sequence 3, Appl	C 227	-17	0.3	2360	1	US-08-456-108-1	Sequence 1, Appl
155	-17	0.3	1112	4	US-09-716-161A-10	Sequence 10, Appl	C 228	-17	0.3	2360	1	US-08-453-703-1	Sequence 1, Appl
156	-17	0.3	1137	4	US-09-543-681A-432	Sequence 432, App	C 229	-17	0.3	2360	1	US-08-453-703-1	Sequence 1, Appl
157	-17	0.3	1155	4	US-09-491-577-21	Sequence 21, Appl	C 230	-17	0.3	2360	1	US-08-456-106-1	Sequence 1, Appl
158	-17	0.3	1178	2	US-08-107-676-26	Sequence 26, Appl	C 231	-17	0.3	2360	2	US-08-456-106-1	Sequence 1, Appl
159	-17	0.3	1178	4	US-09-295-820-26	Sequence 26, Appl	C 232	-17	0.3	2360	2	US-08-456-106-1	Sequence 1, Appl
160	-17	0.3	1206	4	US-09-252-991A-5101	Sequence 5101, Ap	C 233	-17	0.3	2360	4	US-09-633-739-1	Sequence 1, Appl
161	-17	0.3	1211	2	US-08-107-676-2	Sequence 2, Appl	C 234	-17	0.3	2360	5	US-09-633-739-1	Sequence 1, Appl
162	-17	0.3	1211	4	US-09-295-820-2	Sequence 2, Appl	C 235	-17	0.3	2493	5	PCT-US93-03614-1	Sequence 46, App
163	-17	0.3	1212	1	US-08-282-581-1	Sequence 1, Appl	C 236	-17	0.3	2493	5	US-09-976-994-96	Sequence 5, Appl
164	-17	0.3	1212	1	US-08-550-544-1	Sequence 1, Appl	C 237	-17	0.3	2493	1	US-07-977-434-5	Sequence 5, Appl
165	-17	0.3	1218	1	US-09-540-236-1822	Sequence 1822, Ap	C 238	-17	0.3	2493	5	PCT-US91-07035-5	Sequence 93, App
166	-17	0.3	1224	4	US-09-489-039A-3700	Sequence 3700, Ap	C 239	-17	0.3	2540	3	US-08-838-151A-18	Sequence 18, Appl
167	-17	0.3	1231	4	US-09-397-787-34	Sequence 34, Appl	C 240	-17	0.3	2609	4	US-09-976-584-443	Sequence 1, Appl
168	-17	0.3	1245	4	US-09-023-655-1412	Sequence 1412, Ap	C 241	-17	0.3	2619	1	US-08-154-019-3	Sequence 3, Appl
169	-17	0.3	1263	4	US-09-252-991A-15423	Sequence 15423, A	C 242	-17	0.3	2619	1	US-08-461-333-3	Sequence 3, Appl
170	-17	0.3	1265	4	US-09-016-434-1430	Sequence 1430, Ap	C 243	-17	0.3	2619	1	US-08-464-167-3	Sequence 3, Appl
171	-17	0.3	1281	4	US-09-134-000C-996	Sequence 996, App	C 244	-17	0.3	2619	3	US-09-158-313-3	Sequence 3, Appl
172	-17	0.3	1305	4	US-09-252-991A-15557	Sequence 15527, A	C 245	-17	0.3	2619	3	US-08-476-798-3	Sequence 3, Appl
173	-17	0.3	1337	4	US-09-491-577-109	Sequence 109, App	C 246	-17	0.3	2619	3	US-08-476-798-3	Sequence 3, Appl

247	17	0.3	2619	4	US-09-377-497-14	Sequence 14, Appl	320	16	0.3	131	3	US-09-099-639-62	Sequence 62, Appl
248	17	0.3	2686	4	US-09-795-691-1	Sequence 1, Appl	321	16	0.3	131	5	PCT-US93-12588-62	Sequence 62, Appl
C 249	17	0.3	3015	4	US-09-489-039A-2544	Sequence 2544, Ap	322	16	0.3	119	5	PCT-US95-08071-62	Sequence 62, Appl
C 250	17	0.3	3024	2	US-08-836-943-1	Sequence 1, Appl	C 323	16	0.3	157	4	US-09-621-976-15276	Sequence 15276, A
C 251	17	0.3	3034	4	US-09-976-594-636	Sequence 636, App	C 324	16	0.3	155	4	US-09-621-976-15275	Sequence 15275, A
C 252	17	0.3	3060	4	US-09-252-991A-7846	Sequence 7846, Ap	C 325	16	0.3	157	4	US-09-621-976-15278	Sequence 15278, A
253	17	0.3	3366	4	US-09-345-650-2	Sequence 2, Appl	326	16	0.3	182	4	US-09-313-294A-4709	Sequence 4709, Ap
254	17	0.3	3358	4	US-09-223-685-603	Sequence 603, App	327	16	0.3	186	4	US-09-705-448-5	Sequence 5, Appl
255	17	0.3	3458	1	US-08-454-097-34	Sequence 34, Appl	C 328	16	0.3	210	4	US-09-134-001C-2749	Sequence 2749, Ap
256	17	0.3	3505	1	US-08-468-036-43	Sequence 43, Appl	C 329	16	0.3	225	4	US-09-621-976-15270	Sequence 15270, A
257	17	0.3	3505	2	US-08-376-843-43	Sequence 43, Appl	C 330	16	0.3	224	4	US-09-621-976-15277	Sequence 15277, A
258	17	0.3	3505	3	US-08-185-359-34	Sequence 34, Appl	C 331	16	0.3	226	4	US-09-621-976-15272	Sequence 15272, A
259	17	0.3	3601	3	US-09-017-631-23	Sequence 23, Appl	332	16	0.3	227	4	US-09-705-448-6	Sequence 6, Appl
260	17	0.3	3602	2	US-08-883-795A-33	Sequence 33, Appl	333	16	0.3	229	4	US-09-823-381-18	Sequence 18, Appl
261	17	0.3	3602	3	US-09-018-138-1	Sequence 1, Appl	C 334	16	0.3	247	4	US-09-016-434-436	Sequence 436, Appl
262	17	0.3	3602	3	US-09-023-655-1420	Sequence 1420, Ap	C 335	16	0.3	248	4	US-09-016-434-436	Sequence 968, App
263	17	0.3	3900	4	US-09-016-434-1199	Sequence 1199, Ap	C 336	16	0.3	258	4	US-09-252-991A-7356	Sequence 7356, App
264	17	0.3	3984	4	US-09-552-950-2	Sequence 2, Appl	337	16	0.3	270	4	US-08-332-294A-6796	Sequence 5174, Ap
265	17	0.3	4307	4	US-08-915-868-1	Sequence 1, Appl	338	16	0.3	280	4	US-08-313-294A-6796	Sequence 6796, Ap
C 266	17	0.3	4348	2	US-09-566-921-134	Sequence 134, App	339	16	0.3	283	2	US-09-382-552-56	Sequence 2, Appl
C 267	17	0.3	4509	4	US-09-019-172-61	Sequence 61, Appl	340	16	0.3	289	4	US-09-382-552-56	Sequence 56, Appl
C 268	17	0.3	4603	4	US-09-453-702B-22	Sequence 22, Appl	341	16	0.3	292	4	US-09-621-976-8492	Sequence 8492, Ap
269	17	0.3	4632	4	US-09-421-017B-846	Sequence 846, App	342	16	0.3	315	4	US-09-313-294A-5804	Sequence 5804, Ap
270	17	0.3	5124	4	US-09-534-638-2	Sequence 964, App	343	16	0.3	354	4	US-09-489-039A-2664	Sequence 2664, Ap
271	17	0.3	5129	4	US-09-976-594-964	Sequence 1, Appl	C 344	16	0.3	360	4	US-09-621-976-8492	Sequence 2422, Ap
272	17	0.3	5330	4	US-09-023-905A-1	Sequence 5, Appl	C 345	16	0.3	364	4	US-09-337-787-143	Sequence 143, App
273	17	0.3	5889	1	US-08-463-092B-5	Sequence 5, Appl	346	16	0.3	384	4	US-09-134-000C-1544	Sequence 1544, App
274	17	0.3	5889	2	US-08-460-907B-5	Sequence 5, Appl	347	16	0.3	396	4	US-09-640-173-147	Sequence 147, App
275	17	0.3	5889	3	US-08-463-179A-5	Sequence 5, Appl	C 348	16	0.3	411	4	US-09-489-039A-712	Sequence 712, App
276	17	0.3	5889	3	US-08-461-384B-5	Sequence 15, Appl	C 350	16	0.3	414	4	US-09-621-976-11049	Sequence 11049, A
277	17	0.3	6002	1	US-08-698-551-15	Sequence 15, Appl	C 351	16	0.3	420	3	US-09-532-803-E	Sequence 5, Appl
278	17	0.3	6002	2	US-08-602-228-15	Sequence 15, Appl	C 352	16	0.3	420	3	US-09-252-991A-756	Sequence 756, App
279	17	0.3	6002	2	US-08-839-032A-15	Sequence 15, Appl	C 353	16	0.3	421	3	US-09-261-631-12	Sequence 12, Appl
280	17	0.3	6002	2	US-08-185-258C-15	Sequence 15, Appl	C 354	16	0.3	421	3	US-09-265-628-12	Sequence 12, Appl
281	17	0.3	6002	4	US-08-961-527-60	Sequence 60, App	C 355	16	0.3	421	3	US-09-001-141-10	Sequence 10, Appl
C 282	17	0.3	6827	4	US-09-636-215-705	Sequence 705, App	C 356	16	0.3	421	3	US-09-653-403-13	Sequence 13, Appl
C 283	17	0.3	6976	4	US-09-685-166A-705	Sequence 5, Appl	C 357	16	0.3	421	4	US-09-644-597-321	Sequence 321, Appl
C 284	17	0.3	7063	4	US-09-485-632B-5	Sequence 9, Appl	C 358	16	0.3	421	4	US-10-013-784-13	Sequence 321, App
C 285	17	0.3	7812	4	US-09-485-632B-9	Sequence 54, Appl	C 359	16	0.3	421	4	US-09-480-884A-321	Sequence 321, App
C 286	17	0.3	8266	4	US-09-059-584-54	Sequence 5, Appl	C 361	16	0.3	421	4	US-09-542-615A-321	Sequence 321, App
C 287	17	0.3	8266	4	US-09-552-950-5	Sequence 3, Appl	C 362	16	0.3	421	4	US-09-606-421B-321	Sequence 11, Appl
288	17	0.3	9772	4	US-09-312-762A-3	Sequence 3, Appl	C 363	16	0.3	421	4	US-09-786-606-11	Sequence 8508, Ap
289	17	0.3	14707	1	US-08-125-468-1	Sequence 1, Appl	C 364	16	0.3	423	4	US-09-621-976-8508	Sequence 11873, A
C 290	17	0.3	30001	1	US-08-474-933-1	Sequence 1, Appl	C 365	16	0.3	423	4	US-09-489-039A-382	Sequence 382, App
C 291	17	0.3	34001	2	US-09-596-002-18	Sequence 18, Appl	366	16	0.3	423	4	US-09-489-039A-1140	Sequence 1140, Ap
C 292	17	0.3	47981	4	US-09-679-279-1	Sequence 30, Appl	C 367	16	0.3	423	4	US-09-976-594-165	Sequence 18939, A
C 293	17	0.3	58909	4	US-09-596-002-30	Sequence 3, Appl	368	16	0.3	423	4	US-09-621-976-18939	Sequence 18230, A
C 294	17	0.3	75395	4	US-09-984-890-3	Sequence 1, Appl	369	16	0.3	431	4	US-09-621-976-18230	Sequence 15600, A
295	17	0.3	77536	4	US-09-410-551B-1	Sequence 1, Appl	370	16	0.3	431	4	US-09-666-751-59	Sequence 59, Appl
296	17	0.3	98844	4	US-09-791-211-70	Sequence 32, Appl	371	16	0.3	431	4	US-09-621-976-18938	Sequence 18938, A
297	17	0.3	118067	4	US-09-497-855A-32	Sequence 40, Appl	372	16	0.3	431	4	US-09-621-976-18939	Sequence 17757, A
C 298	17	0.3	161652	4	US-09-497-855A-40	Sequence 17, Appl	373	16	0.3	431	4	US-09-621-976-18939	Sequence 3838, Ap
C 299	17	0.3	169998	4	US-09-676-610B-24	Sequence 24, Appl	374	16	0.3	431	4	US-09-621-976-18939	Sequence 174, App
C 300	17	0.3	176373	3	US-09-128-155-17	Sequence 10, Appl	375	16	0.3	431	4	US-09-621-976-18939	Sequence 893, App
C 301	17	0.3	197496	4	US-09-877-177A-10	Sequence 1, Appl	376	16	0.3	431	4	US-09-621-976-18939	Sequence 564, App
C 302	17	0.3	580073	4	US-08-545-588D-1	Sequence 3, Appl	377	16	0.3	431	4	US-09-621-976-18939	Sequence 9, Appl
C 303	17	0.3	786431	4	US-09-751-389-3	Sequence 16, Appl	378	16	0.3	431	4	US-09-621-976-18939	Sequence 565, App
C 304	17	0.3	24	4	US-09-653-754-3	Sequence 16, Appl	379	16	0.3	431	4	US-09-621-976-18939	Sequence 153, App
C 305	16	0.3	25	4	US-09-599-287A-16	Sequence 168, App	380	16	0.3	431	4	US-09-621-976-18939	Sequence 653, App
C 306	16	0.3	65	4	US-09-363-939A-164	Sequence 36, Appl	381	16	0.3	431	4	US-09-621-976-18939	Sequence 1541, Ap
C 307	16	0.3	65	4	US-09-363-939A-168	Sequence 36, Appl	382	16	0.3	431	4	US-09-621-976-18939	Sequence 3245, App
C 308	16	0.3	78	3	US-08-479-744A-36	Sequence 36, Appl	383	16	0.3	431	4	US-09-621-976-18939	Sequence 1541, Ap
C 309	16	0.3	78	3	US-08-280-757B-36	Sequence 5, Appl	384	16	0.3	431	4	US-09-621-976-18939	Sequence 8727, App
310	16	0.3	78	3	US-09-227-585-5	Sequence 5, Appl	385	16	0.3	431	4	US-09-621-976-18939	Sequence 1827, App
311	16	0.3	78	3	US-09-425-762-96	Sequence 9, Appl	386	16	0.3	431	4	US-09-621-976-18939	Sequence 8727, App
312	16	0.3	80	3	US-09-133-321-9	Sequence 88, Appl	387	16	0.3	431	4	US-09-621-976-18939	Sequence 1827, App
313	16	0.3	129	1	US-09-209-676-88	Sequence 62, Appl	388	16	0.3	431	4	US-09-621-976-18939	Sequence 8727, App
314	16	0.3	131	1	US-07-998-003A-62	Sequence 62, Appl	389	16	0.3	431	4	US-09-621-976-18939	Sequence 8727, App
315	16	0.3	131	1	US-08-453-274B-62	Sequence 62, Appl	390	16	0.3	431	4	US-09-621-976-18939	Sequence 8727, App
316	16	0.3	131	1	US-08-453-274B-62	Sequence 62, Appl	391	16	0.3	431	4	US-09-621-976-18939	Sequence 8727, App
317	16	0.3	131	1	US-08-453-274B-62	Sequence 62, Appl	392	16	0.3	431	4	US-09-621-976-18939	Sequence 8727, App
318	16	0.3	131	1	US-08-268-161A-62	Sequence 62, Appl							
319	16	0.3	131	2	US-08-453-702A-62	Sequence 62, Appl							

C 393	16	0.3	552	4	US-09-489-039A-3494	Sequence 3494, Ap	C 466	16	0.3	954	4	US-10-043-238-2	Sequence 2, Appl1
C 394	16	0.3	571	4	US-09-404-879A-48	Sequence 48, Appl	C 467	16	0.3	963	4	US-09-071-035-285	Sequence 285, Ap
C 395	16	0.3	571	4	US-09-338-933-48	Sequence 48, Appl	C 468	16	0.3	963	4	US-09-134-000C-4132	Sequence 2133, Ap
C 396	16	0.3	571	4	US-09-215-681-48	Sequence 48, Appl	C 469	16	0.3	966	4	US-09-252-991A-1439	Sequence 7433, Ap
C 397	16	0.3	571	4	US-09-216-003A-48	Sequence 48, Appl	C 470	16	0.3	972	4	US-09-252-991A-1956	Sequence 2956, Ap
C 398	16	0.3	576	1	US-08-086-428B-2	Sequence 2, Appl1	C 471	16	0.3	987	4	US-09-489-039A-1674	Sequence 11517, A
C 399	16	0.3	576	2	US-08-468-570-2	Sequence 2, Appl1	C 472	16	0.3	990	4	US-09-252-991A-11517	Sequence 270, Ap
C 400	16	0.3	576	2	US-08-290-665A-2	Sequence 2, Appl1	C 473	16	0.3	1001	4	US-09-641-638-270	Sequence 446, Ap
C 401	16	0.3	576	4	US-08-466-601A-2	Sequence 2, Appl1	C 474	16	0.3	1001	4	US-09-641-638-446	Sequence 290, Ap
C 402	16	0.3	576	4	PCT-US95-10398-2	Sequence 2, Appl1	C 475	16	0.3	1001	4	US-09-641-638-446	Sequence 106, Ap
C 403	16	0.3	579	4	US-09-252-991A-8637	Sequence 8637, Ap	C 476	16	0.3	1007	4	US-09-328-477C-106	Sequence 106, Ap
C 404	16	0.3	582	4	US-09-621-976-18934	Sequence 18934, A	C 477	16	0.3	1011	4	US-09-641-806-3	Sequence 3, Appl1
C 405	16	0.3	603	3	US-09-385-988-233	Sequence 233, Ap	C 478	16	0.3	1011	4	US-09-723-123-3	Sequence 3, Appl1
C 406	16	0.3	605	4	US-09-643-597-207	Sequence 207, Ap	C 479	16	0.3	1011	4	US-09-252-991A-14782	Sequence 14782, A
C 407	16	0.3	605	4	US-09-480-884A-207	Sequence 207, Ap	C 480	16	0.3	1011	4	US-09-722-862-3	Sequence 3, Appl1
C 408	16	0.3	605	4	US-09-543-615A-207	Sequence 207, Ap	C 481	16	0.3	1026	4	US-09-252-991A-5229	Sequence 5229, Ap
C 409	16	0.3	605	4	US-09-606-421B-207	Sequence 207, Ap	C 482	16	0.3	1026	4	US-09-641-806-1	Sequence 1, Appl1
C 410	16	0.3	615	1	US-08-539-304A-1	Sequence 1, Appl1	C 483	16	0.3	1026	4	US-09-723-123-1	Sequence 1, Appl1
C 411	16	0.3	615	1	US-09-134-001C-763	Sequence 763, Ap	C 484	16	0.3	1026	4	US-09-722-862-1	Sequence 1, Appl1
C 412	16	0.3	630	4	US-09-252-991A-656	Sequence 656, Ap	C 485	16	0.3	1026	4	US-09-252-991A-15460	Sequence 15460, A
C 413	16	0.3	631	3	US-08-545-809A-31	Sequence 31, Appl	C 486	16	0.3	1041	4	US-09-489-039A-194	Sequence 194, Appl
C 414	16	0.3	631	3	US-08-943-731-187	Sequence 187, Ap	C 487	16	0.3	1044	1	US-07-975-528-1	Sequence 1, Appl1
C 415	16	0.3	651	4	US-09-489-039A-957	Sequence 957, Ap	C 488	16	0.3	1044	4	US-07-974-409C-422	Sequence 422, Ap
C 416	16	0.3	654	4	US-09-252-991A-11814	Sequence 11814, A	C 489	16	0.3	1044	4	US-07-974-409C-422	Sequence 422, Ap
C 417	16	0.3	669	4	US-09-221-017B-61	Sequence 61, Appl	C 490	16	0.3	1047	4	US-09-252-991A-1034	Sequence 9034, Ap
C 418	16	0.3	672	4	US-09-489-039A-6264	Sequence 6264, Ap	C 491	16	0.3	1065	3	US-09-325-897-1	Sequence 1, Appl1
C 419	16	0.3	688	3	US-09-328-111-453	Sequence 453, Ap	C 492	16	0.3	1065	4	US-09-837-726-1	Sequence 1, Appl1
C 420	16	0.3	696	1	US-07-656-566-4	Sequence 4, Appl1	C 493	16	0.3	1071	4	US-09-543-681A-3985	Sequence 3985, Ap
C 421	16	0.3	697	6	5171843-10	Patent No. 5171843	C 494	16	0.3	1086	3	US-09-280-409-3	Sequence 3, Appl1
C 422	16	0.3	701	3	US-09-331-625A-55	Sequence 55, Appl	C 495	16	0.3	1088	4	US-09-755-665-61	Sequence 61, Appl
C 423	16	0.3	701	3	US-09-133-321-1	Sequence 1, Appl1	C 496	16	0.3	1090	4	US-09-755-665-61	Sequence 61, Appl
C 424	16	0.3	701	4	US-09-494-151-55	Sequence 55, Appl	C 497	16	0.3	1090	4	US-09-755-665-60	Sequence 60, Appl
C 425	16	0.3	701	4	US-09-972-484-55	Sequence 55, Appl	C 498	16	0.3	1091	4	US-09-755-665-60	Sequence 60, Appl
C 426	16	0.3	701	4	US-09-621-976-206	Sequence 206, Ap	C 499	16	0.3	1098	4	US-09-755-665-62	Sequence 62, Appl
C 427	16	0.3	703	4	US-09-023-655-1097	Sequence 1097, Ap	C 500	16	0.3	1101	4	US-09-424-712-1	Sequence 1, Appl1
C 428	16	0.3	710	3	US-08-998-416-603	Sequence 603, Ap	C 501	16	0.3	1103	4	US-09-424-712-2	Sequence 2, Appl1
C 429	16	0.3	715	4	US-09-247-155-139	Sequence 139, Ap	C 502	16	0.3	1110	4	US-09-543-681A-3342	Sequence 3342, Ap
C 430	16	0.3	726	4	US-09-621-976-1875	Sequence 1875, Ap	C 503	16	0.3	1123	4	US-09-755-665-15	Sequence 15, Appl
C 431	16	0.3	729	4	US-09-252-991A-11877	Sequence 11877, A	C 504	16	0.3	1137	6	5171843-8	Patent No. 5171843
C 432	16	0.3	730	3	US-09-352-990-3	Sequence 3, Appl1	C 505	16	0.3	1139	4	US-09-690-454-13	Sequence 13, Appl
C 433	16	0.3	738	2	US-08-604-989A-8	Sequence 8, Appl1	C 506	16	0.3	1140	4	US-09-252-991A-14999	Sequence 14999, A
C 434	16	0.3	741	4	US-09-149-476-202	Sequence 202, Ap	C 507	16	0.3	1140	4	US-09-328-352-4089	Sequence 4089, Ap
C 435	16	0.3	741	4	US-09-252-991A-12066	Sequence 12066, A	C 508	16	0.3	1149	4	US-09-252-991A-14783	Sequence 14783, A
C 436	16	0.3	741	4	US-09-489-039A-5512	Sequence 5512, Ap	C 509	16	0.3	1176	4	US-09-543-681A-3917	Sequence 3917, Ap
C 437	16	0.3	742	2	US-08-966-316-2	Sequence 2, Appl1	C 510	16	0.3	1179	2	US-09-489-039A-5553	Sequence 5553, Ap
C 438	16	0.3	765	3	US-08-331-625A-40	Sequence 40, Appl	C 511	16	0.3	1179	2	US-08-465-794-4	Sequence 4, Appl1
C 439	16	0.3	765	4	US-09-494-151-40	Sequence 40, Appl	C 512	16	0.3	1182	1	US-09-049-812-4	Sequence 4, Appl1
C 440	16	0.3	765	4	US-09-252-991A-7969	Sequence 7969, Ap	C 513	16	0.3	1182	1	US-07-640-476-4	Sequence 4, Appl1
C 441	16	0.3	768	4	US-09-972-484-40	Sequence 40, Appl	C 514	16	0.3	1182	1	US-09-252-991A-6534	Sequence 6534, Ap
C 442	16	0.3	768	4	US-09-252-991A-9221	Sequence 9221, Ap	C 515	16	0.3	1185	6	5290690-1	Patent No. 5290690
C 443	16	0.3	780	4	US-09-651-656-26	Sequence 26, Appl	C 516	16	0.3	1197	4	US-09-252-991A-10197	Sequence 10197, A
C 444	16	0.3	780	4	US-09-650-855-26	Sequence 26, Appl	C 517	16	0.3	1197	4	US-09-691-641-43	Sequence 43, Appl
C 445	16	0.3	783	4	US-09-252-991A-11569	Sequence 11569, A	C 518	16	0.3	1203	4	US-09-540-236-1433	Sequence 1433, Ap
C 446	16	0.3	793	4	US-09-252-991A-3048	Sequence 3048, Ap	C 519	16	0.3	1206	4	US-09-489-039A-823	Sequence 823, Ap
C 447	16	0.3	813	4	US-09-252-991A-12621	Sequence 12621, A	C 520	16	0.3	1212	4	US-09-489-039A-823	Sequence 823, Ap
C 448	16	0.3	826	3	US-09-071-035-287	Sequence 287, Ap	C 521	16	0.3	1215	4	US-09-352-991A-16442	Sequence 16442, A
C 449	16	0.3	835	3	US-08-957-302A-9	Sequence 9, Appl1	C 522	16	0.3	1215	4	US-09-543-681A-3574	Sequence 3574, Ap
C 450	16	0.3	835	3	US-09-542-403-9	Sequence 9, Appl1	C 523	16	0.3	1226	4	US-09-328-352-1841	Sequence 1841, Ap
C 451	16	0.3	836	3	US-08-998-416-332	Sequence 332, Ap	C 524	16	0.3	1226	4	US-09-540-236-1436	Sequence 1436, Ap
C 452	16	0.3	840	4	US-09-252-991A-10769	Sequence 10769, A	C 525	16	0.3	1245	1	US-09-252-991A-1749	Sequence 1749, A
C 453	16	0.3	840	4	US-09-252-991A-14967	Sequence 14967, A	C 526	16	0.3	1245	1	US-07-887-072B-1	Sequence 1, Appl1
C 454	16	0.3	843	4	US-09-252-991A-2744	Sequence 2744, Ap	C 527	16	0.3	1245	1	US-08-466-444-1	Sequence 1, Appl1
C 455	16	0.3	858	4	US-09-172-952-8	Sequence 8, Appl1	C 528	16	0.3	1248	4	US-09-252-991A-6118	Sequence 6118, Ap
C 456	16	0.3	870	4	US-09-489-039A-764	Sequence 764, Ap	C 529	16	0.3	1248	4	US-09-252-991A-9872	Sequence 9872, Ap
C 457	16	0.3	876	4	US-09-513-783A-9	Sequence 9, Appl1	C 530	16	0.3	1251	4	US-09-540-236-295	Sequence 295, Ap
C 458	16	0.3	876	4	US-09-252-991A-8946	Sequence 8946, Ap	C 531	16	0.3	1252	4	US-09-673-395A-73	Sequence 73, Appl
C 459	16	0.3	876	4	US-09-252-991A-14476	Sequence 14476, A	C 532	16	0.3	1257	4	US-09-489-039A-6981	Sequence 6981, Ap
C 460	16	0.3	891	4	US-09-489-039A-360	Sequence 360, Ap	C 533	16	0.3	1284	3	US-08-331-625A-42	Sequence 42, Appl
C 461	16	0.3	897	5	PCT-US86-0530A-303	Sequence 303, Ap	C 534	16	0.3	1284	4	US-09-494-151-42	Sequence 42, Appl
C 462	16	0.3	915	3	US-08-557-614-1	Sequence 1, Appl1	C 535	16	0.3	1284	4	US-09-972-484-42	Sequence 42, Appl
C 463	16	0.3	915	3	US-09-252-991A-14968	Sequence 14968, A	C 536	16	0.3	1287	4	US-09-489-039A-394	Sequence 394, Ap
C 464	16	0.3	918	4	US-09-252-991A-12009	Sequence 12009, A	C 537	16	0.3	1299	1	US-08-723-938-4	Sequence 4, Appl1
C 465	16	0.3	933	4	US-09-252-991A-8606	Sequence 8606, Ap	C 538	16	0.3	1299	2	US-09-080-538-4	Sequence 4, Appl1